SEGMENTATION OF SUBSTANTIA NIGRA AND RED NUCLEUS IN MRI IMAGES: APPLICATION TO PARKINSON'S DISEASE

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Computer Science

at the

University of Canterbury

by

Dibash Basukala

March 2021

This thesis is dedicated to my beloved parents,

Mr. Ganga Pd. Basukala and Mrs. Shree Maya Basukala

ABSTRACT

Accurate segmentation of substantia nigra (SN) and red nucleus (RN) is challenging, yet important for understanding health problems like Parkinson's disease (PD). This thesis proposes robust algorithms to segment SN and RN from T2*-weighted images and quantitative susceptibility mapping (QSM) MRI. After optimising the algorithms to produce reliable SN and RN segmentations, multiple MRI metrics extracted from the SN and RN are used to investigate association with cognitive and motor function in PD.

For T2*-weighted images, an adaptive *k*-means algorithm and a combination of level set method and wavelet transform are proposed for accurate segmentation of substantia nigra. The latter method is further improved with substantia nigra seeding, contrast enhancement and watershed transform to provide efficient segmentations of substantia nigra and red nucleus in QSM images. The algorithm-derived segments of SN and RN are compared to expert manually derived ('gold standard') segmentations using the Dice similarity index for quantitative assessment.

The QSM values extracted from manually drawn SN and RN are compared to values extracted from the proposed segmentations in a subsample of 40 participants. Additionally, QSM values obtained from both 'gold standard' and proposed segmentation method are compared between groups (PD and healthy controls), and the relationship with global cognitive ability and motor severity in PD are investigated using Bayesian regression models.

The proposed segmentation algorithms produce effective segmentations of SN and RN from T2*-weighted images and QSM images, validated against expert manual segmentation. This thesis shows moderate evidence of increased QSM values in SN in PD relative to healthy controls, with moderate evidence for association between QSM, global cognitive ability, and motor impairment in the SN in PD. RN QSM values also exhibit moderate evidence for association with motor impairments in PD.

After validating the proposed segmentation algorithm, SN and RN are segmented in a large cohort of 127 PD participants and 47 matched controls. Multiple MRI metrics are extracted from the SN and RN regions from different modalities and include QSM, R2*, diffusion tensor imaging (DTI) parameters, perfusion, and volume. The MRI metrics are investigated independently to assess their relationship with group (PD vs

controls), global cognitive ability, and motor impairment, again using Bayesian regression models.

This thesis shows strong evidence of increased QSM, R2*, MD and RD values, and moderate evidence of decreased FA values in SN, moderate evidence of increased QSM, R2* and MD values in RN in PD relative to controls. SN MD values also demonstrate moderate evidence of association with motor impairments in PD.

Lastly, group status (PD/Control) is predicted based on multiple MRI metrics from the SN and RN. Four different classifiers are compared with 10-fold cross validation: linear support vector machine (SVM), non-linear SVM, logistic regression, and random forest.

In terms of predicting disease group, linear SVM gave the highest performance with area under the receiver operating characteristics curve (AUCROC) = 0.76 and area under the precision-recall curve (AUCPR) = 0.89, showing robust classification of disease based on MRI values from SN and RN.

When applying the proposed robust segmentation method to a cohort of healthy controls and PD patients, this thesis demonstrates interesting association between groups and clinical scores utilizing various MRI measurements from substantia nigra and red nucleus, suggesting that the combination of different MRI metrics may provide useful information about iron deposition, microstructural alterations, perfusion changes, and volume atrophy of the midbrain nuclei in Parkinson's disease, and also may be helpful in distinguishing PD from controls. Therefore, we suggest that an efficient and improved midbrain segmentation algorithm may be utilized to monitor disease severity in Parkinson's. Additionally, the segmentation method proposed in this thesis may also be applied to different scenarios and disease groups where SN and RN are of particular interest.

Keywords—Region segmentation, substantia nigra, red nucleus, quantitative susceptibility mapping, Parkinson's disease, k-means algorithm, thresholding, substantia nigra seeding, contrast enhancement, level set method, wavelet transform, watershed transform, Bayesian regression, iron deposition, global cognitive ability, motor impairments, support vector machine.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Professor Ramakrishnan Mukundan from the core of my heart for his expert guidance and knowledge throughout my PhD degree. His unwavering support and encouragement were valuable during my study. I appreciate his motivation, patience, kindness, and advice throughout my research journey. I am also grateful to my co-supervisor Dr. Tracy R. Melzer who expertly guided me throughout this dissertation. His constant support, motivation, and feedback were extremely valuable for my research.

I would also like to acknowledge Dr. Ross J. Keenan, Dr. Anthony Lim, and Dr. Michael A. Hurrell for their expert advice and suggestions from the clinical perspective. I am thankful to the aforementioned radiologists' for the subjective evaluation of the outputs produced in this research. The clinicians also provided the manual tracings of the structures, important for this research.

I would also like to thank Dr. Daniel J. Myall and Dr. Reza Shoorangiz, for their valuable help and inputs in the data analysis and machine learning aspects of this research. I appreciate Professor Tim J. Anderson, clinical director of the New Zealand Brain Research Institute, Christchurch, New Zealand, and Professor John C. Dalrymple-Alford for all the resources, important for this research. I am also grateful to the entire team of the New Zealand Brain Research Institute, directly or indirectly involved in this research.

I would like to take time to acknowledge the University of Canterbury Doctoral Scholarship, without which I would not have been able to pursue my doctoral studies.

I would also like to thank all the technical and administrative staff of the Department of Computer Science and Software Engineering for all the help and support extended during my dissertation. I also appreciate the friendly and pleasant workspace of the department.

I am thankful to the many current and former colleagues and friends including Tieta, Alan, Prerna, Ori, Hai Peng, David, Anthony, Sarmad, Simon, Julio, Josh, and Oliver for all the joyous moments and many afternoon coffee chats.

I want to thank all my friends I have had over the years, particularly Lynn and Shajeel, for always backing me up. Finally, I want to express my gratitude to my family, whose value to me only grows with age. I know I do not say as much, but I love you all.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AD	Axial Diffusivity
ANIMAL	Automatic Nonlinear Image Matching and Anatomical Labelling
ASL	Arterial Spin Labelling
A-STAR	Anchor-Based Street and Traffic Aware Routing
AUCPR	Area under the Precision-Recall Curve
AUCROC	Area under the Receiver Operating Characteristics Curve
BET	Brain Extraction Tool
CBF	Cerebral Blood Flow
CI	Confidence Interval
CN	Caudate Nucleus
CSF	Cerebrospinal Fluid
CV	Chan and Vese
DARTEL	Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra
DN	Dentate Nucleus
DSC-MRI	Dynamic Susceptibility Contrast enhanced MRI
DT-CWT	Dual-Tree Complex Wavelet Transform
DTI	Diffusion Tensor Imaging
DWT	Discrete Wavelet Transform
FA	Fractional Anisotropy
FIRST	FMRIB Integrated Registration and Segmentation Tool
FLAIR	Fluid Attenuated Inversion Recovery
FN	Fascicula Nigrale
FOV	Field of View
FW	Free Water
GAS	Geodesic Active Surface

GM	Gray Matter
GP	Globus Pallidus
GPe	External Globus Pallidus
GPi	Internal Globus Pallidus
GRE	Gradient Echo
HAMA	Hamilton Anxiety Scale
ICC	Intraclass Correlation Coefficient
ICV	Intracranial Volume
LEDD	Levodopa Equivalent Daily Dosage
MADRS	Montgomery Asberg Depression Rating Scale
MD	Mean Diffusivity
MDS	Movement Disorders Society
MEDI	Morphology Enabled Dipole Inversion
MNI	Montreal Neurological Institute
MoCA	Montreal Cognitive Assessment
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
NMS-MRI	Neuromelanin Sensitive MRI
NZBRI	New Zealand Brain Research Institute
PCA	Principal Component Analysis
PD	Parkinson's Disease
PDD	Parkinson's Disease Dementia
PDE	Partial Differential Equation
PD-MCI	Parkinson's Disease with Mild Cognitive Impairment
PD-N	Parkinson's Disease with Normal Cognition
PPB	Parts per Billion
PPM	Parts per Million

PR	Precision-Recall
PUT	Putamen
QSM	Quantitative Susceptibility Mapping
RD	Radial Diffusivity
RN	Red Nucleus
ROC	Receiver Operating Characteristics
ROI	Region of Interest
RSF	Region-Scalable Fitting
SD	Standard Deviation
SN	Substantia Nigra
SNPC	Substantia Nigra Pars Compacta
SNPR	Substantia Nigra Pars Reticulata
SPGR	Spoiled Gradient Recalled Echo
SPM	Statistical Parametric Mapping
STN	Subthalamic Nucleus
SVM	Support Vector Machine
SWI	Susceptibility-Weighted Images
TE	Echo Times
THA	Thalamus
TI	Inversion Time
TR	Repetition Time
UPDRS-III	Unified Parkinson's Disease Rating Scale Part 3
VBA	Voxel Based Analysis
VBM	Voxel-Based Morphometry
WM	White Matter
3T	3 Tesla
7T	7 Tesla

1 INTRODUCTION

Parkinson's disease (PD) is a progressive neurodegenerative disorder that affects over 4 million people worldwide, which is expected to reach 9.3 million by 2030 (Dorsey et al., 2007). PD involves the loss of dopaminergic neurons in the substantia nigra (SN). This cell death affects purposeful movement and eventually results in loss of motor coordination (Brain & Walton, 1977). Usually, by the time of diagnosis, approximately 60 – 80 percent of dopaminergic neurons within the SN would have been lost (Fearnley & Lees, 1991). In addition to cell death, the SN also experiences increased iron accumulation (Martin, Wieler, & Gee, 2008; Oakley et al., 2007; Ward, Zucca, Duyn, Crichton, & Zecca, 2014), which may play a role in the pathogenesis of PD. Though its exact role remains unclear, iron accumulation may encourage oxidative stress, alpha synuclein aggregation, and the formation of Lewy pathology (Bjorklund, Hofer, Nurchi, & Aaseth, 2019; Munoz, Carrasco, Campos, Aguirre, & Nunez, 2016; Wan et al., 2017).

The red nucleus (RN) is a pair of small gray matter structures which lie adjacent to the substantia nigra in the midbrain. It is also an iron rich structure (Drayer et al., 1986) and is thought to be involved in motor coordination and cognitive functions. RN infarction leads to tremor, asynergia, intellectual fatigability, decreased verbal fluency and discrete memory impairment.

1.1 Research Motivation

Accurate and precise identification of the structures is key to understanding the role that SN- and RN-specific processes play in PD progression such as iron deposition,

microstructural alterations, perfusion deficits, and volume atrophy which can be quantified using multiple magnetic resonance imaging (MRI) modalities and will be described in detail in the next section. However, segmentation of these nuclei is challenging because of their small size, unclear boundaries, morphometric variability and similar intensity profiles with the adjacent structures on standard T1- and, to a lesser extent, T2-weighted magnetic resonance imaging. Furthermore, tremor is one of the defining characteristics of PD, and therefore motion on all types of brain imaging can confound identification of these, and other, structures. Quantitative susceptibility mapping (QSM), a novel MRI technique is capable of producing high-resolution images of these smaller midbrain structures (and other subcortical structures) with excellent image contrast reflecting their intrinsic levels of iron content. Previous work has implemented expert manual tracing of SN (Acosta-Cabronero et al., 2017), but accurate automated segmentation is preferable to manual segmentation because of the absence of operator bias, the potential for greater consistency, and application to large cohorts in a systematic way. However, there is only a limited literature available for automated segmentation of these midbrain structures (Garzón, Sitnikov, Bäckman, & Kalpouzos, 2018; Guo et al., 2018; Haegelen et al., 2013; Kim, Lenglet, Duchin, Sapiro, & Harel, 2014; Visser, Keuken, Forstmann, & Jenkinson, 2016; Yiming Xiao et al., 2012; Y. Xiao et al., 2014), because high quality imaging capable of visualizing these smaller structures has only recently emerged.

1.2 Magnetic Resonance Imaging

A number of different magnetic resonance imaging modalities were utilized in this thesis. These included T1-weighted, quantitative susceptibility mapping, R2*, diffusion tensor imaging (DTI), and arterial spin labelling (ASL) MRI. Note, all these images were acquired within the same one-hour session and can be routinely acquired on the majority of clinical 3T MRI scanners. We also used T2*-weighted MRI obtained from an online database (Forstmann et al., 2014b) as an additional resource.

1.2.1 T1-weighted Imaging

The T1-weighted spoiled gradient recalled echo (SPGR) image is a high-resolution anatomical image used for structural analysis in this thesis. It also serves as the base image for all subsequent image processing: that is, all other images are coregistered to the high resolution SPGR prior to further preprocessing. The SPGR image provides 1

mm isotropic resolution with good contrast between gray and white matter and cerebral spinal fluid. Figure 1.1 presents an example of a T1-weighted SPGR image.

Figure 1.1: An axial slice of a T1-weighted SPGR image.

1.2.2 T2-weighted Imaging*

T2*-weighted imaging is acquired with a gradient echo (GRE) sequence. The GRE sequence of T2*-weighted imaging requires high uniformity of the magnetic field and hence can detect even very slight changes, helping to highlight subtle tissue changes (Chavhan, Babyn, Thomas, Shroff, & Haacke, 2009). Figure 1.2 presents an example of a T2*-weighted image.

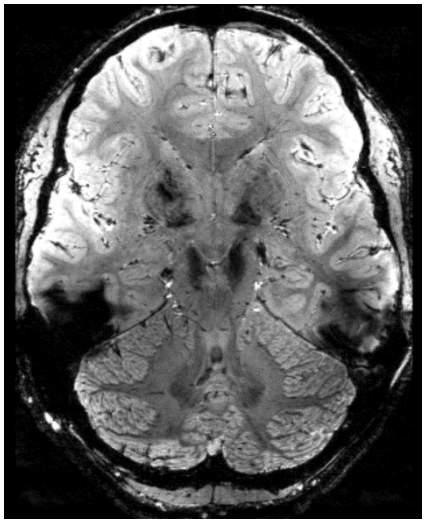


Figure 1.2: An axial slice of a T2*-weighted image (acquired from an online database (Forstmann et al., 2014b)).

1.2.3 Quantitative Susceptibility Mapping

Quantitative susceptibility mapping (Pelzer, Florin, & Schnitzler, 2019) is a relatively new MRI technique for quantifying magnetic susceptibility (Haacke et al., 2015). Essentially, QSM solves the inverse problem from a magnetic field to susceptibility source and thus maps the magnetic property of local tissue (de Rochefort et al., 2010). Magnetic susceptibility χ, is defined as the degree of magnetisation of a material (δM) subject to a unit magnetic field (δH) and is an intrinsic property of all materials,

$$\chi = \frac{\delta M}{\delta H} \qquad \textbf{1.1}$$

M and *H* are both magnetic fields with the unit Ampere per metre. Hence, χ is dimensionless.

Magnetic susceptibility varies depending on the chemical constituents of the material and their bonding type and can either be positive (magnetisation parallel with the external field or paramagnetic) or negative (magnetisation anti-parallel with the field or diamagnetic). Therefore χ is a complex value with both magnitude and direction. This property of χ causes the observed χ value to vary depending on the shape and orientation of the object being investigated relative to the externally applied magnetic field (Schenck, 1996).

Iron (stored as ferritin) is ferromagnetic and has one of the highest magnetic susceptibilities of any element, 200,000 (SI units); even in salt solution it has a χ, 0.00075 SI units, that is roughly 2 orders of magnitude higher than human soft tissue, -0.000009 (SI units) (Schenck, 1996). Therefore, small differences in iron accumulation in certain basal ganglia structures may be detectable with QSM and could act as a non-invasive marker of Parkinson's disease.

From the same acquisition, R2* (which is defined as 1/T2*) may also be determined. R2* is the effective transverse relaxation rate, another property of tissue. Iron deposition induces local magnetic field inhomogeneities which lead to faster signal decay and thus to an increase in relaxation rate (Yablonskiy & Haacke, 1994). Therefore, R2* may also provide a measure of pathological iron accumulation in vivo. However, R2* mapping depends on field strength (Yao et al., 2009), contains substantial blooming artifacts that increase with echo time (S. Wang et al., 2013), and generally relates to iron

concentration in a complex way (C. Liu, Li, Tong, Yeom, & Kuzminski, 2015). Therefore, this thesis will investigate variations in iron concentration using both QSM and R2* images. Figure 1.3 presents an example of a QSM image and a R2* image.

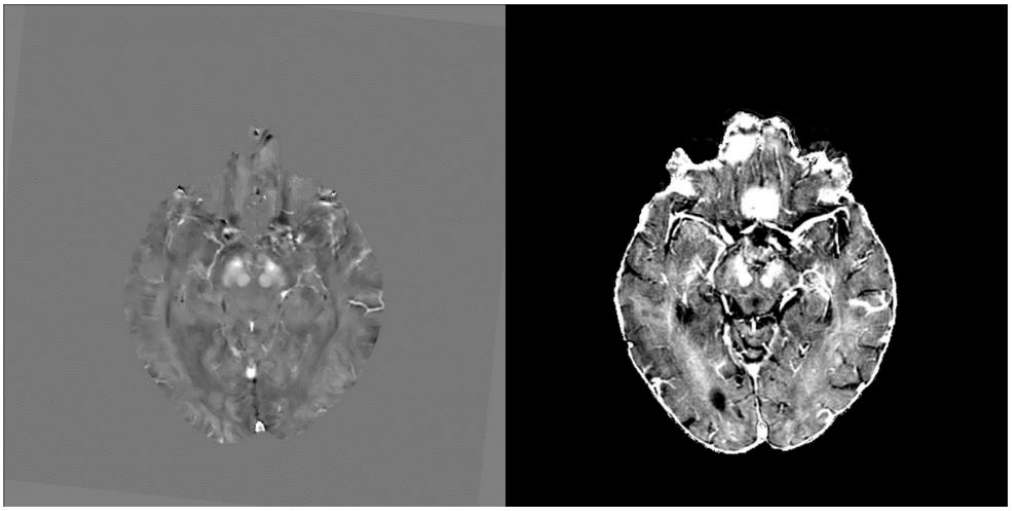


Figure 1.3: An axial slice of a QSM image (left panel) and R2* image (right panel).

1.2.4 Diffusion Tensor Imaging

Diffusion MRI is a non-invasive technique that is uniquely sensitive to tissue microstructure. The MR signal can be made sensitive to the microscopic motion of water molecules. Because the microscopic motion of water is influenced by the geometrical structure of the environment, this MR contrast can be used to probe tissue microstructure. Diffusion MRI allows quantitative measures of the magnitude of motion of water and its directional variation, and can therefore provide unique information about microstructural changes in the brain (Basser, 1995; Johansen-Berg & Behrens, 2006).

Properties of water diffusion in the brain are described as being either isotropic or anisotropic. In tissues with isotropic diffusion, water molecules diffuse equally in all directions, such as in the ventricles. Anisotropic diffusion refers to the case where water diffuses preferentially along a specific direction. This occurs in areas such as white matter tracts, where tightly bunched, myelinated fibres cause water to diffuse anisotropically, preferentially diffusing along, and not perpendicular to, the fibres. The diffusion tensor is the most common model used to describe the directional dependence of water diffusion in the brain (Le Bihan et al., 2001). The tensor can be represented as an ellipsoid. The surface of the ellipsoid represents how water diffuses within a voxel of

tissue where the quantities λ_1, λ_2 and λ_3 represent the major and minor axes of the ellipsoid, and do not depend on its orientation (Figure 1.4).

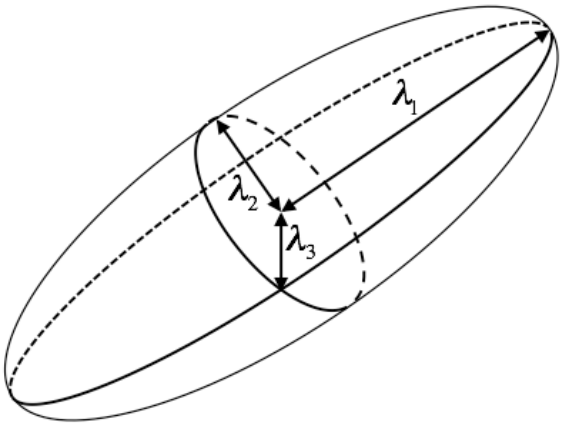


Figure 1.4: Schematic of the diffusion tensor ellipsoid. λ_1, λ_2 and λ_3 are the eigenvalues of the diffusion tensor and correspond to the major and minor axes of the ellipsoid. The direction of each eigenvalue (or ellipsoid axis) is the corresponding eigenvector (ε_1, ε_2 and ε_3). The long axis, or principal eigenvector, indicates the direction of largest diffusion and predominant fibre orientation. Adapted from Watts et al. (Watts, Liston, Niogi, & Uluğ, 2003).

A wide range of direction-independent parameters can be calculated from the diffusion tensor (Figure 1.5) (Kingsley, 2006). Mean diffusivity (MD) is a measure of the average diffusion over all directions or the magnitude of diffusion. It is the average of the three eigenvalues, measured in mm^2/s. Fractional anisotropy (FA) is used to quantify the degree of angular variation (anisotropy). FA values vary from 0 to 1, where 0 is indicative of isotropic diffusion (the diffusion ellipsoid is a sphere) and 1 is where diffusion is constrained along a single direction (the ellipsoid appears long and thin). Axial diffusivity (AD) is the diffusion along the direction of greatest diffusion (λ_1), while radial diffusivity (RD) quantifies diffusion along the minor axes of the ellipsoid (average of λ_2 and λ_3). In general (and simplistically), a decreased in FA and an increase in MD, AD, and RD is interpreted as reflecting damage/loss of microstructural integrity.

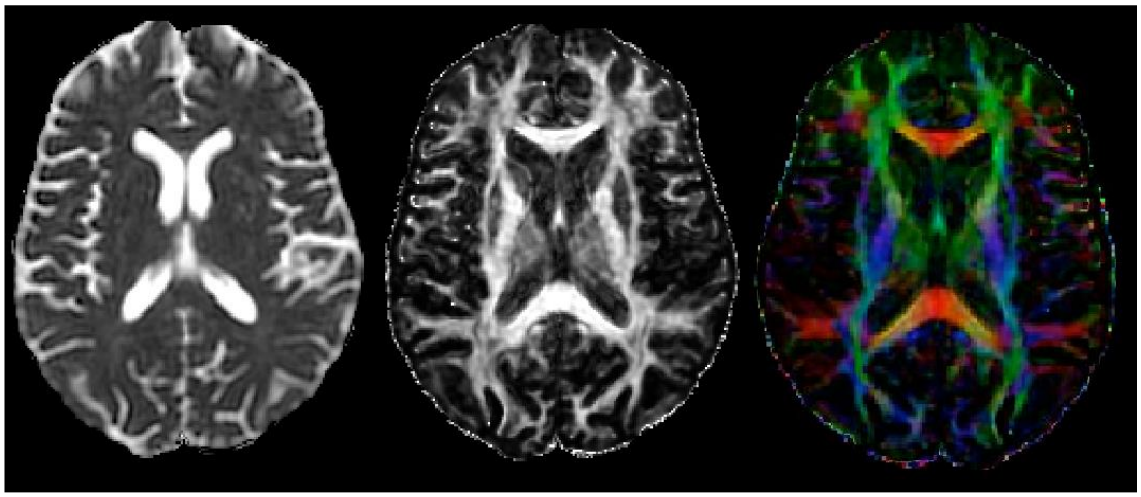


Figure 1.5: An example of an MD image (left), FA image (middle), and FA colour map image (right) derived from diffusion tensor imaging. Red indicates white matter fibres in the left-right orientation; green, anterior-posterior; and blue, inferior-superior.

1.2.5 Arterial Spin Labelling

Arterial spin labelling is a non-invasive MRI method that produces absolute quantification of cerebral perfusion (Detre, Leigh, Williams, & Koretsky, 1992). Perfusion refers to the capillary blood supply delivering oxygen and nutrients to brain tissue, measured in $ml.min^{-1}.100g^{-1}$ (Detre, Wang, Wang, & Rao, 2009). Although slower and with a lower signal-to-noise ratio than dynamic susceptibility contrast enhanced (DSC)-MRI, ASL is completely non-invasive (and therefore excludes the gadolinium contrast which may, rarely, cause a severe allergic reaction) and can be repeated as often as necessary. In relation to radiotracer techniques, ASL eliminates expensive, potentially harmful radioactive materials; long preparation times, and requires shorter scan times. ASL is an easily repeatable addition to routine MRI scanning and produces absolute perfusion images. Thus, ASL provides an attractive method to assess cerebral perfusion non-invasively, both in health and disease.

Instead of an exogenous tracer (e.g. ^{15}O or gadolinium), ASL uses magnetically labelled arterial water as an endogenous tracer. This magnetic label is typically produced by saturating or inverting the longitudinal component of the MR signal of the in-flowing protons. Once reaching the capillary bed, the labelled or tagged water passes into the brain tissue reducing the local tissue's longitudinal magnetization (Aguirre, Detre, & Wang, 2005). The greater the flow into the imaging slice, the larger the reduction in signal. In this thesis, the arterial water was inverted using pseudo-

continuous labelling (Dai, Garcia, de Bazelaire, & Alsop, 2008). After a specified delay time in which the tagged blood is allowed to flow into the imaging volume (1.5 seconds in this thesis), an image is acquired. This image is known as the tagged or labelled image. A control image is then acquired of the same imaging volume using the same sequence, but this time the inflowing blood is not tagged. The difference between the un-tagged and tagged images is used to quantify perfusion (cerebral blood flow (CBF)) (Alsop et al., 2015). Figure 1.6 displays an example of an ASL-acquired perfusion image.

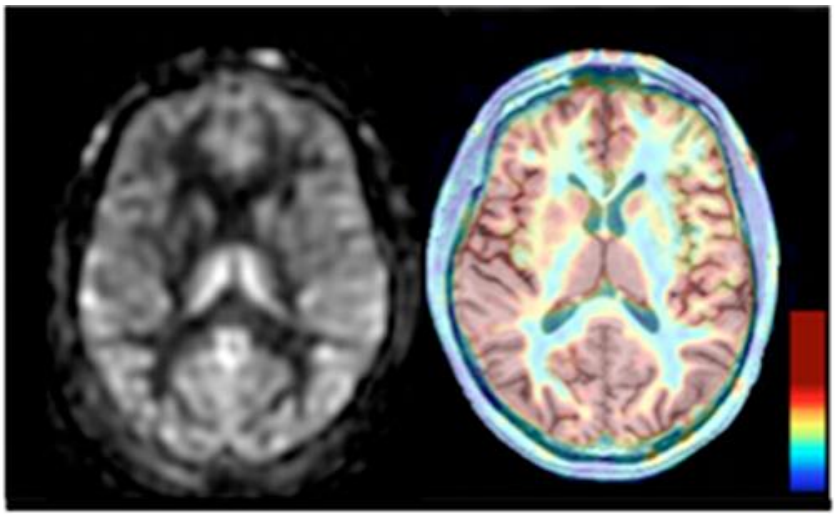


Figure 1.6: An example of a quantitative, ASL-acquired perfusion image. Left panel: the raw perfusion image. Right panel: the perfusion image overlaid on a high-resolution structural image. The colour bar runs from 0 (blue) to 100 (red) ml/100g/min.

ASL does suffer from a number of disadvantages. The technique is associated with inherently low signal-to-noise of the signal difference between the labelled and control images. Many signal averages are required to overcome this limitation. These numerous averages, in addition to the necessity of acquiring both a tagged and un-tagged image to create a single quantitative perfusion image, considerably extend imaging time. A delay is required between labelling and imaging to allow the labelled blood to reach the imaging plane, but this necessary delay reduces the magnitude of the label and further increases scanning times. Despite these limitations, ASL provides a unique, non-invasive opportunity to investigate easily acquired, absolute perfusion images in PD.

1.3 Research Objectives

As stated in the earlier section, there exists only a few techniques for the segmentation of smaller midbrain structures. Therefore, the main objective of this research is to develop accurate, effective, and robust algorithms for the substantia nigra and red nucleus segmentation, specifically from T2*-weighted images and QSM images. The goal of this research is also to evaluate the segmentation outputs from the proposed methods through the subjective assessment performed by the radiologists' and the quantitative analysis based on the manual segmentation drawn by clinical neuroradiology experts'.

The thesis will compare algorithm-derived segmentations of SN and RN in T2*-weighted and QSM images to 'gold standard' segments manually identified by experienced radiologists'. The thesis will then investigate the relationship between iron accumulation and PD, specifically examining the difference in QSM values between PD and healthy controls, and association with cognitive and motor dysfunction in PD.

The thesis will then implement the validated algorithm and produce automated segmentations of the midbrain structures using QSM images obtained from a large cohort of Parkinson's disease and healthy controls to investigate multi-modal MRI metrics in SN and RN associated with PD and cognitive impairment in PD. This will be followed by an attempt to classify this large group of participants as PD or control using machine learning (linear support vector machine (SVM)) with multi-modal MRI, including R2*, diffusion tensor imaging parameters, cerebral blood flow (quantified using ASL), volume, and QSM of the SN and RN across groups. Assessing multiple modalities may help us to understand iron deposition, microstructural and perfusion changes, and volume atrophy occurring in PD patients and also to develop the best classification of PD/control based on multiple MRI metrics from the SN and RN.

1.4 Research Methodology

The research methodology can be broadly classified into three main stages in accordance to the research goals set prior, which are described below.

Stage 1: In this stage, the algorithms for the substantia nigra segmentation using T2*-weighted images is developed. Two different segmentation algorithms based on weighted thresholding approach, and the combination of level set method and wavelet transform are proposed. Subjective (qualitative) assessment is performed by

collaborating radiologists' to evaluate the results. Furthermore, Dice score (quantitative assessment) is estimated based on the manual segmentations of SN performed by the clinical expert for the analysis of the segmentation outputs.

Stage 2: In this stage, the algorithm that is robust and performing close to gold standard segmentation from the previous stage is selected; precisely methodology based on the level set method is adopted for further improvements. Extensions added to the algorithm include substantia nigra seeding, contrast enhancement and watershed transform, which improve segmentation and separation of substantia nigra and red nucleus from QSM images. The results are analysed both by qualitative and quantitative assessments.

The QSM values are extracted from the proposed automated method and gold standard SN and RN segments (right and left are extracted independently). Correlation between the values extracted from the two methods are assessed using the Pearson correlation coefficient and intraclass correlation coefficient (ICC). Furthermore, QSM values from both methods are compared across groups (healthy controls and PD), global cognitive ability, and motor functions using Bayesian regression models.

Stage 3: In the final stage, the proposed segmentation method is applied to QSM images from a large group of healthy controls and PD participants, after obtaining consistent results from the proposed method and gold standard segmentations (across groups and clinical scores). The QSM values along with R2*, DTI metrics (FA, AD, MD, RD), CBF, and volume extracted from both substantia nigra and red nucleus segments are compared across groups (both PD/control and across cognitive subgroups), global cognitive ability, and motor impairments in Parkinson's disease. Furthermore, predictive power is assessed to distinguish healthy controls and PD using linear SVM classifier.

1.5 Thesis Contributions

1. An improved algorithm for the segmentation of substantia nigra using T2*-weighted image is proposed. The methodology employs weighted thresholding approach based on *k*-means algorithm (D. Basukala, Mukundan, Melzer, & Keenan, 2018b). This proposed method results in an effective segmentation of SN, validated using qualitative and quantitative assessment that will be discussed in detail in Chapter 4.

2. A second improved algorithm for the substantia nigra segmentation using level set method and wavelet transform (D. Basukala, Mukundan, Melzer, & Keenan, 2018a) is proposed utilizing T2*-weighted image. The SN segments obtained using this method is found to be superior in comparison to the previous method, and will be fully described in Chapter 4.

3. The algorithm based on level set method and wavelet transform is further improved for the automated segmentation of substantia nigra and red nucleus based on QSM images (D. Basukala, Mukundan, Melzer, & Lim, 2019). The enhanced methodology overall includes substantia nigra seeding, contrast enhancement, level set method, wavelet transform, and watershed transform steps resulting in high quality segmentations of SN and RN, validated against expert manual segmentation in 40 participants. QSM values extracted from SN and RN segments are used to investigate the variations of iron in Parkinson's disease, and moderate evidence of associations are found in SN which will be described in detail in Chapter 5 (Dibash Basukala et al., 2021).

4. Automated segmentation of substantia nigra and red nucleus is further performed using QSM images obtained from a large cohort of healthy controls and PD (n = 174). Multi-modal MRI metrics are extracted from automated SN and RN segmentations, including QSM, R2*, DTI parameters, CBF, and volume. The different MRI metrics are assessed to compare across groups (both PD/control and across cognitive subgroups), cognitive and motor impairments in PD, and interesting evidence of associations is obtained. Linear SVM classifier utilizing multi-modal MRI features also results in good accuracy in separating PD and controls, which will be presented in detail in Chapter 6.

1.6 List of Research Publications

1. Basukala, D., Mukundan, R., Melzer, T., & Keenan, R. (2018). Segmentation of Substantia Nigra Using Weighted Thresholding Method. *International Conference on Image and Vision Computing New Zealand (IVCNZ),* Auckland, New Zealand, 19-21 Nov 2018. Doi: 10.1109/ivcnz.2018.8634679.

2. Basukala, D., Mukundan, R., Melzer, T., & Keenan, R. (2018). Segmentation of Substantia Nigra for the Automated Characterization of Parkinson's Disease. *IEEE International Conference on Image Processing, Applications and Systems (IPAS),* Sophia Antipolis, France, 12-14 Dec 2018, pp. 85-90. Doi: 10.1109/ipas.2018.8708886.

3. Basukala, D., Mukundan, R., Melzer, T., & Lim, A. (2019). Automated Segmentation of Substantia Nigra and Red Nucleus in Quantitative Susceptibility Mapping Images. *20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT),* Gold Coast, Australia, 5-7 Dec 2019, pp. 378-383. Doi: 10.1109/pdcat46702.2019.00074. (*This paper was selected for publication of an extended version in Computers and Electrical Engineering Journal (Elsevier)*)

4. Basukala, D., Mukundan, R., Lim, A., Hurrell, M. A., Keenan, R. J., Dalrymple-Alford, J. C., Anderson, T. J., Myall, D. J., & Melzer, T. R. (2021). Automated Segmentation of Substantia Nigra and Red Nucleus Using Quantitative Susceptibility Mapping Images: Application to Parkinson's Disease. *Computers and Electrical Engineering,* 91, 107091. Doi: 10.1016/j.compeleceng.2021.107091.

1.7 Thesis Organization

This section gives an overview of the contents of each chapter.

Chapter 2 describes the literature review, which establishes the background to the research work carried out in the thesis. Different segmentation techniques proposed for the midbrain structures are described and reviewed followed by the quantification of brain iron deposition in Parkinson's disease. It also provides a review of different MRI

measurements used in the field and their contribution to the characterization and separation of PD from controls.

Chapter 3 describes the materials and methods used for the research presented in the thesis. This part of the thesis describes two different databases used for the experimental work. Additionally, it also describes the metric used for the assessment of the segmentation accuracy as well as the correlation coefficients.

Chapter 4 describes two different substantia nigra segmentation algorithms proposed in the thesis using T2*-weighted images viz. (i) segmentation based on weighted thresholding, and (ii) segmentation based on level set method and wavelet transform. Experimental results for both methods are presented along with the assessment of accuracy of segmentation.

Chapter 5 describes the improved segmentation method based on level set and wavelet transform with additional substantia nigra seeding, contrast enhancement, and watershed transform. The methodology is applied for the segmentation of substantia nigra and red nucleus from QSM images. Experimental results and Dice scores are presented, and discussed in detail. QSM values compared across groups and clinical scores are demonstrated for both proposed method and gold standard segmentations. The analysis in this and next chapter are performed with the inputs and suggestions from statistical expert Dr. Daniel J. Myall from the New Zealand Brain Research Institute.

Chapter 6 describes the multi-modal assessment of QSM, R2*, DTI parameters, CBF and volume from SN and RN segments across groups (both PD/control and across cognitive subgroups), global cognitive ability, and motor impairments. The SN and RN segments are obtained from the proposed method described in the previous chapter. It also presents the strength of combination of different MRI measurements to separate healthy controls and PD using linear SVM. The machine learning aspects of this chapter are performed with the inputs and suggestions from Dr. Reza Shoorangiz from the New Zealand Brain Research Institute.

Chapter 7 concludes the thesis and outlines future direction of research in the area of segmentation and characterization of brain diseases utilizing multi-modal MRI.

2 LITERATURE REVIEW

Medical image segmentation (Grau, Mewes, Alcaniz, Kikinis, & Warfield, 2004; Weisenfeld, Mewes, & Warfield, 2006) is an essential task in many medical image analysis applications including surgical planning, three-dimensional visualization, post-surgical assessment, volumetric measurement, radiotherapy planning and detection of changes over time. In medical image segmentation, magnetic resonance imaging is used because of its ability to generate tomographic images with high spatial resolution and contrast. Moreover, MRI is considered as a non-invasive imaging modality as MRI requires no ionizing radiation for image acquisition. The segmentation of key brain structures from MRI has received great attention as it plays an important role in understanding brain changes associated with a number of different diseases and disorders, including Parkinson's disease, dementia, Alzheimer's disease, multiple sclerosis, epilepsy, and cerebral atrophy. This thesis will particularly focus on the segmentation problem that is applicable to Parkinson's disease. Accurate and robust segmentation of the midbrain structures including substantia nigra and red nucleus can help understand Parkinson's disease and explore its relationship with iron content and microstructural integrity across the disease spectrum. This chapter gives a detailed discussion of the relevant literature for the segmentation of the midbrain nuclei, quantification of brain iron deposition in PD and combined assessment of different MRI measurements in PD. This chapter aims to fulfil the following objectives.

- To provide a description of the current literature and how others have approached segmentation of the substantia nigra and red nucleus.

- To highlight the motivation behind specific methodological steps utilized in the thesis.
- To identify gaps in the research and limitations of the methods currently used in the literature.

2.1 Segmentation Techniques

This section focuses on the review of existing segmentation methods proposed for midbrain structures. There exists limited literature for automated segmentation of the midbrain nuclei, potentially because high quality and high contrast imaging capable of visualizing these smaller structures is a recent development. All the segmentation methods reviewed in this section of thesis falls under the category of structural segmentation methods i.e., they are either atlas, template, or prior shape based.

Xiao et al. (Yiming Xiao et al., 2012) proposed an atlas-based segmentation of the subthalamic nucleus (STN), red nucleus and substantia nigra for deep brain stimulation by incorporating multiple MRI contrasts. T1-weighted image contrast was inadequate for visualizing the deep brain nuclei like STN, RN and SN. Therefore, T1-weighted and T2*-weighted images were combined to achieve a better contrast for the visualization of these structures. The combined contrast helped improve non-rigid registration to the atlas and eventually higher segmentation accuracy was achieved compared to T1-weighted – T1-weighted registration. The proposed method was validated on 6 subjects which included 3 healthy and 3 PD patients. *This thesis performs contrast enhancement by using local gray level information and global gray level information of the image. In this thesis, only a single imaging modality (T2*-weighted or quantitative susceptibility mapping) is used at a time in the process.*

Haegelen et al. (Haegelen et al., 2013) proposed the automated segmentation of basal ganglia and deep brain structures in MRI of Parkinson's disease patients. Manual segmentation was compared with three automated methods including automatic nonlinear image matching and anatomical labeling (ANIMAL), symmetric image normalization (SyN) and patch-label fusion method. ANIMAL and SyN are registration-based methods performed on T1-weighted images of PD patients, while the patch-based method was performed on T2-weighted images of PD patients. ANIMAL and SyN used a single labeled template; therefore, it was susceptible to inter-subject variability and was time-consuming. On the other hand, the patch-based technique used

fusion of several manually labeled images; hence, it was better at handling inter-subject variability. Moreover, this technique was much faster than the registration-based methods. However, all three methods resulted in similar mean and median kappa values, which meant they produced equivalent segmentation results of deep brain structures. Hence, the patch-based method was equally good and faster than the frequently used template-based technique.

Xiao et al. (Y. Xiao et al., 2014) performed an investigation of morphometric variability of subthalamic nucleus, red nucleus, and substantia nigra in advanced Parkinson's disease patients. Label-fusion method was used for the segmentation of STN, RN and SN in 3T resolution T2-weighted MRI of advanced PD patients. Label-fusion segmentation was performed using majority-voting approach for the segmentation of the midbrain nuclei. Anatomical variability was estimated using the Talairach atlas. Moreover, principal component analysis (PCA) was used to measure the morphometric variability, as PCA gives a more detailed and intuitive description of the variability of these deep brain nuclei than the conventional coordinate analysis. The authors reported that there exists a correlation between principal components and Unified Parkinson's Disease Rating Scale part 3 (UPDRS-III) scores, which illustrates a relationship between a measure of motor impairment and morphometric variability.

Kim et al. (Kim et al., 2014) proposed a semiautomatic segmentation of brain subcortical structures from high-field MRI. The proposed method used susceptibility-weighted images (SWI), T2-weighted images and fractional anisotropy images obtained from diffusion MRI at 7T. The geodesic active surface (GAS) model failed to perform accurate segmentation of subcortical structures because of low contrast, weak boundaries, and occlusion around objects. Hence, the GAS model was improved by integrating shape prior information, a new multimodal edge indicator function and anatomical constraints between adjacent structures. The shape prior information was obtained by manually segmenting other subjects or on a different scan date if the same subject was considered. The new multimodal edge indicator function was obtained by the combination of two edge maps acquired from the Laplacian of SWI, FA, or T2-weighted images with boundary information obtained from the shape priors. The anatomical constraints between adjacent structures were performed by using a penalty term to avoid overlapping between the adjacent structures. The authors claim that the proposed approach enhanced the segmentation quality of subcortical structures.

Visser et al. (Visser et al., 2016) proposed an automatic segmentation of the substantia nigra, subthalamic nucleus, and red nucleus in 7T data at young and old age. T2*-weighted images and Quantitative Susceptibility Mapping images were used for the automated segmentation. The proposed method consisted of two main components an intensity model and a shape model. Appearance of edges close to the border of the structure to be segmented was described by an intensity model, while the deformation of the reference shape was described by a shape model. The shape model was based on a Markov random field (MRF) and it helped in smoothing the segmentation. The proposed method required a single reference mesh and a set of unlabelled training volumes for the training stage. The authors claim that the proposed automatic method resulted in the segmentation that corresponded well with manual delineations. The automated segmentation was performed on the dataset (Forstmann et al., 2014a) consisting of scans of healthy participants from different age groups including young, middle-aged and elderly. This was one of the major advantages of the proposed method, as most of the state-of-the-art algorithms used only older PD patients. *This thesis employs wavelet-based smoothing using dual-tree complex wavelet transform (DT-CWT) to remove the jaggedness present in the segmentation outputs. Moreover, the data used in this research consists of both healthy controls and Parkinson's disease patients.*

Garzón et al. (Garzón et al., 2018) proposed an automated segmentation of the midbrain structures (substantia nigra, red nucleus, and subthalamic nucleus) with high iron content. The study was performed on QSM, R2* and T2-weighted fluid attenuated inversion recovery (FLAIR) images obtained from 40 healthy subjects consisting of both younger and older individuals. The algorithm used multi-atlas estimation of spatial priors for the structures by averaging a set of non-linearly registered manual labels to the midbrain. The spatial priors were then used for a Gaussian mixture model to classify the image voxels imposing smoothness constraints to ensure anatomical plausibility. The authors claim that the proposed method effectively segments the midbrain structures and also reported that the volumes estimated from both manual and automated segmentations were significantly correlated. The algorithm showed higher accuracy in QSM images compared to R2* and FLAIR images. It has also been reported that the susceptibility values were significantly higher in older individuals compared to younger in both manual and automated segmentations. However, the main drawback of the proposed methodology was the need to perform multiple non-linear registrations as

well as its sensitivity to the number of atlases. *This thesis extends the study to find the correlation between quantitative susceptibility mapping values of healthy controls and Parkinson's disease patients. Moreover, correlation between susceptibility values and clinical scores including global cognitive scores and UPDRS-III scores in PD patients is also presented as cognitive deficits and motor impairment have been constantly reported in PD patients.*

Guo et al. (Guo et al., 2018) proposed the seed point discontinuity-based segmentation method for the substantia nigra and red nucleus in quantitative susceptibility mapping images. The data was obtained in a 3T system and the study was performed in 22 subjects, which included 11 PD patients, and 11 healthy subjects. The proposed method used atlas registration to locate the two seed points in SN and RN. The anchor-based street and traffic aware routing (A-STAR) algorithm was employed to find the shortest spatial path between two seed points. An intensity profile of the pixels was used along the shortest path to remove the voxels, which did not belong to SN or RN. The seed point discontinuity implies that if there is any connection between two seed points, they must be removed at the border of the SN and RN so that the level set method would work effectively in separating them. The authors claim that the proposed method effectively segments SN and RN and was better than the level set method, single-atlas method and multi-atlas method. However, the proposed method was still dependant on atlas registration for finding the seed points which is an important limitation of this method. The authors stated that the accuracy of the proposed method decreased considerably if the atlas failed to find the seed points. *In contrast, this thesis uses substantia nigra seeding to produce seeds within left and right substantia nigra to initiate the segmentations. The level set method has been used for the segmentation of substantia nigra and red nucleus in this research work.*

Currently, deep learning methods are widely used for the segmentation of brain MRI (Akkus, Galimzianova, Hoogi, Rubin, & Erickson, 2017; Ding et al., 2020; Grøvik et al., 2021; Lundervold & Lundervold, 2019; Milletari et al., 2017; Park et al., 2019; Sendra-Balcells et al., 2020) with superlative accuracy, and experts' claiming it to be the future of research. However, we focus on the segmentation based on classical methods because of several drawbacks of deep learning including the requirement of large datasets and annotations for training and validation. Deep learning based models result in poor approximation when the training data is limited; which is common in brain MRI particularly focusing on midbrain and deep brain structures. Deep learning

architectures also require expert manual ground truth labels which is tedious on a large-scale data. Similarly, the popular deep learning methods are well known for their high computational complexity.

2.2 Quantifying Brain Iron Deposition in Parkinson's disease

Iron is consistently found to be increased in the midbrain region of Parkinson's disease patients relative to healthy individuals. Therefore, this section aims to discuss existing literature available for the quantification of brain iron deposition and explore its relationship across groups and clinical scores.

Barbosa et al. (Barbosa et al., 2015) quantified brain iron deposition in patients with Parkinson's disease using quantitative susceptibility mapping, R2 and R2*. The study was conducted from data obtained from 30 healthy controls (64 $\pm$7 years old) and 20 idiopathic PD patients (66 $\pm$ 8 years old). The segmentation of substantia nigra, substantia nigra pars compacta (SNpc), red nucleus, globus pallidus (GP), putamen (PUT), caudate nucleus (CN), thalamus (THA), white matter (WM) and gray matter (GM) was conducted manually. An unpaired t-test was used to compare the iron concentration between the two groups. A significant increase in QSM, R2 and R2* values in both SN and SNpc were observed for the PD patients except for R2* values in SN, confirming the increase of iron in PD patients. However, the highest significant difference was found in the SNpc using QSM. Moreover, it was also reported that the QSM had higher sensitivity to classify PD patients for both SN and SNpc than the R2 and R2* maps. The study was conducted in the small number of patients, which is the main limitation of this study. Similarly, the evaluation was done only on the most representative slice of the region of interest (ROI). *In contrast, this thesis uses three-four slices containing the majority of the SN/RN in every subject for the evaluation of iron content in both healthy controls and Parkinson's disease patients with a broader distribution of disease durations.*

Murakami et al. (Murakami et al., 2015) studied the usefulness of quantitative susceptibility mapping for the diagnosis of Parkinson's disease. The study was conducted on the QSM and R2* images obtained from 21 controls and 21 patients with Parkinson's disease. The region of interest including substantia nigra, red nucleus, globus pallidus, caudate nucleus, putamen, and thalamus were manually traced by neuroradiologists using Image J software. The ROIs were analyzed separately for each

imaging technique (one at a time). The authors reported that the QSM and R2* values for the SN were significantly higher in PD patients than in controls but the same was not true for other brain nuclei. However, QSM was more sensitive than R2* in discriminating PD patients and controls. There was no difference in QSM values between the affected side and unaffected side of the brain (PD generally presents with motor symptoms on one side of the body. The contralateral brain hemisphere is referred to as the 'affected side'). Additionally, receiver operating characteristics (ROC) curve analysis revealed that QSM was more accurate in discriminating PD patients from controls than R2* maps. However, the study had several limitations including the limited sample size and the controls not being healthy volunteers. The authors also stated that the ROIs used for QSM and R2* mapping were not exactly the same which might affect the results of the accuracy of QSM and R2* values. *This thesis includes a cohort of controls who are healthy individuals. Moreover, the same region of interest is used for different imaging modalities.*

Du et al. (Du et al., 2016) performed study on quantitative susceptibility mapping of the midbrain in Parkinson's disease. The study was conducted using QSM and R2* maps obtained from 47 PD patients and 47 healthy controls. Voxel based analysis (VBA) and region of interest analysis were performed in normalized space to analyze QSM values and R2* values. VBA was performed using statistical parametric mapping (SPM) version 12 while ROI-based analysis was performed using SAS 9.3. The authors reported that both QSM values and R2* values in the substantia nigra pars compacta were significantly higher in the PD patients than in controls. However, the sensitivity of QSM was higher than R2* for delineating PD-related changes in the SNpc. Similarly, it has also been reported that the QSM values in the SNpc were significantly correlated with clinical measures including disease duration, Levodopa Equivalent Daily Dosage (LEDD), and Unified Parkinson's Disease Rating Scale part II, but not with the R2* values. The authors stated that the cross-sectional nature of the study is its important limitation.

An et al. (An et al., 2018) performed study on quantifying iron deposition within the substantia nigra of Parkinson's disease by quantitative susceptibility mapping. For the purpose of this study, QSM images were obtained from 31 healthy controls and 44 PD patients. Manual segmentation of SN was performed and QSM values extracted. The authors reported that the QSM values in SN is significantly increased in PD patients compared to controls, and the increment is more prominent with advanced disease

severity stage. SN iron content was also reported to be significantly correlated with the clinical measures; PD with mild severity showed correlation with only part of the clinical measures while PD with advanced severity showed significant correlation with all the clinical measures recruited in the study (Hoehn-Yahr stage, UPDRS-III, UPDRS, Montgomery Asberg Depression Rating Scale (MADRS) and Hamilton Anxiety Scale (HAMA) scores). Similarly, it was reported that akinetic/rigidity-dominant PD patients was the most affected subgroup among tremor-dominant PD, akinetic/rigidity-dominant PD and mixed-PD. However, the study was conducted on the relatively small sample size of PD patients, which is its important limitation. Similarly, the authors also reported that in spite of its advantages, QSM signal might still be affected by various factors like calcium, lipid, or myelin content.

Chen et al. (Chen et al., 2019) performed study on iron deposition in Parkinson's disease by quantitative susceptibility mapping. QSM images were obtained from 26 healthy volunteers and 33 PD patients. The region of interest including substantia nigra pars compacta, substantia nigra pars reticulata (SNpr), red nucleus, fascicula nigrale (FN), internal globus pallidus (GPi), external globus pallidus (GPe), subthalamic nucleus, putamen, caudate nucleus, thalamus, dentate nucleus (DN) and cortex were manually drawn. The authors reported that iron content was significantly increased in SNpc, RN, FN, GPi, PUT, CN of PD patients compared to the healthy volunteers. However, only QSM values of SNpc, SNpr and RN were found to be significantly increased in late-stage PD patients compared to the early-stage PD patients. Similarly, it has also been reported that there was no pattern of increased iron deposition along the GP-FN-SN pathway. The ROI was drawn manually by the researchers, which was the significant limitation of the study. *On the other hand, the region of interest used in this thesis is automatically segmented and validated by a total number of three radiologists' having years of experience in neuroimaging.*

Uchida et al. (Uchida et al., 2019) performed study on voxel based quantitative susceptibility mapping in Parkinson's disease with mild cognitive impairment. The study was conducted in 20 healthy controls, 22 Parkinson's disease with normal cognition (PD-N) and 24 Parkinson's disease with mild cognitive impairment (PD-MCI). Voxel-based QSM analysis and region of interest-based analysis were compared between the groups. WFUPickAtlas (fmri.wfubmc.edu/software/pickatlas) software was used to automatically map the ROIs to brain atlas space. The authors reported that PD-MCI had significantly increased magnetic susceptibility across cuneus, precuneus,

caudate head, fusiform gyrus and orbitofrontal cortex compared to PD-N using voxel based QSM analysis. In the PD patients, QSM values were negatively correlated with Montreal Cognitive Assessment (MoCA), a global cognitive ability screening tool, in the cuneus and caudate head while positive correlation was observed with UPDRS-III in the putamen. ROI based analysis revealed that QSM values in PD-N were significantly higher in the substantia nigra, globus pallidus and putamen compared to healthy controls. The authors also stated that the substantia nigra might be too small for voxel-based analysis. However, the study had several limitations including the dependence on Movement Disorders Society (MDS) Task Force Level I criteria for PD-MCI diagnosis and cross-sectional design of the study. *In contrast, the research work in this thesis is based on cognition assessed by comprehensive neuropsychological testing fulfilling the Movement Disorders Society Task Force Level II criteria.*

Thomas et al. (Thomas et al., 2020) performed study on brain iron deposition linked with cognitive severity in Parkinson's disease. The study was performed on 37 controls and 100 PD patients (early-stage PD – mid-stage PD). Whole-brain QSM statistical analysis were performed using SPM12 and Randomise V.2.9 and threshold-free cluster enhancement (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise) in FSL. SPM misclassified substantia nigra as white matter; therefore, SN was manually traced to perform regional QSM statistical analysis. The authors reported that QSM in PD was found to be increased in prefrontal cortex and putamen compared to controls. ROI analysis of SN also revealed higher QSM values in PD than in controls. Similarly, it was also reported that QSM increment with decreasing MoCA scores was observed in the hippocampus, thalamus, and caudate nucleus of PD patients. The authors also further stated that the UPDRS-III score of PD patients was found to be significantly associated with QSM in the putamen, however no significant relationship was observed between QSM and UPDRS-III in SN. In contrast, voxel-based morphometry (VBM) analysis revealed no atrophy difference between controls and PD or any association with clinical scores. The authors suggested that segmentation of the whole-brain results in splitting of areas such as thalamic nuclei, and therefore findings in this region may have been distorted. *The research work in this thesis is conducted with the participants ranging from healthy controls to Parkinson's disease with normal cognition, mild cognitive impairment, and dementia.*

2.3 Combined Assessment of R2*, DTI Parameters, ASL and Volume

Magnetic resonance imaging such as R2*, DTI, ASL (CBF quantification), and volume measure different tissue characteristics, but may provide complementary information. Therefore, this section describes existing literature available for the combined assessment of MR parameters and explore its contribution for the distinction between PD and healthy controls.

Péran et al. (Péran et al., 2010) performed study on magnetic resonance imaging markers of Parkinson's disease nigrostriatal signature. The study was conducted on 30 PD patients and 22 controls. In each subject, thalamus, caudate, putamen and pallidum were automatically segmented on the T1-weighted images with FMRIB integrated registration and segmentation tool (FIRST) 1.2 integrated within the FSL. On the other hand, SN and RN were manually segmented by radiologists on T2*-weighted images. The segmented ROIs thus obtained were coregistered into fractional anisotropy, mean diffusivity, and R2* maps, and the mean values were computed. The authors reported that the R2* values were significantly increased in SN, FA was decreased in SN and thalamus, and MD was increased in thalamus in PD patients compared to controls and the results were confirmed by voxel-based analysis. Moreover, logistic regression was performed to classify PD and controls. The authors also reported that the combination of three markers (R2* in SN, FA in SN, and MD in putamen or caudate nucleus) were sufficient to obtain the discriminant power of >95% while the separation between PD and controls were not as high when the markers were taken individually suggesting that R2*, FA and MD provide different but complementary information. However, the main limitation of the study was the use of limited number of subjects. *This thesis performs the automatic segmentation of substantia nigra and red nucleus in QSM images obtained from large number of subjects (including both healthy controls and PD) for analyzing the group differences, cognitive subgroups, cognitive deficits, and motor impairments. In addition to QSM, R2*, volume, FA, AD, MD, RD and CBF values were also extracted to perform the analysis and eventually to perform classification between healthy controls and PD.*

Du et al. (G. Du et al., 2011) performed study on combined R2* and Diffusion tensor imaging changes in the substantia nigra in Parkinson's disease. The study was conducted in 16 PD and 16 control subjects. The SN was manually segmented on T2-

weighted images using ITK-SNAP (http://www.itksnap.org) (Yushkevich et al., 2006) and mapped onto R2* and FA maps and the means were eventually calculated. The authors reported that the R2* values in SN were significantly increased and FA values were decreased in PD compared to controls while there were no significant difference in RD, MD and AD between PD and controls. Furthermore, there were no significant correlation between R2* and FA suggesting that these two MRI measurements may reflect different pathological processes occurring in the SN of PD. The authors also reported that they did not find correlation of MRI measures with UPDRS-III and disease duration. Additionally, logistic regression revealed that the combination of R2* and FA provided higher discrimination between PD and controls than using the MRI measures independently. However, the limited sample size used in the study was its significant limitation. *This thesis extracted median QSM, R2*, FA, AD, MD, RD, CBF, and approximate volume from substantia nigra and red nucleus segmentations to perform the analysis.*

Loane et al. (Loane et al., 2016) performed study on aberrant nigral diffusion in Parkinson's disease. The study was conducted on diffusion images obtained from 18 early stage PD patients and 14 healthy controls. The PD patients were scanned twice; once at baseline and the other at follow-up (19.3 months apart). The ROIs including substantia nigra, caudate and putamen were manually drawn for the purpose of the study. The authors reported that PD patients demonstrated significantly decreased FA and increased MD in SN compared to controls at follow-up and was also confirmed with longitudinal analysis (PD baseline vs PD follow-up). However, no significant FA or MD differences were observed in the caudate and putamen between controls and PD, and over time within Parkinson's disease patients. The study was conducted on a small number of subjects, which was its important limitation.

Pelizzari et al. (Pelizzari et al., 2019) performed combined assessment of diffusion parameters and cerebral blood flow within basal ganglia in early Parkinson's disease. The study was conducted with 1.5T MRI images obtained from 26 PD patients and 26 healthy controls. Multi-contrast PD25 atlas (Y. Xiao et al., 2015) was used to generate SN, RN and STN while the Harvard-oxford atlas was used for defining caudate, putamen, globus pallidus, and thalamus masks. Median CBF, FA, AD, MD, and RD values were extracted from the ROIs. The authors reported that no significant differences were observed in FA, MD, and RD between PD and healthy controls while increased AD was observed in the putamen in PD. Furthermore, no significant CBF

differences were observed between the groups. There were also no difference in gray matter volume between PD and healthy controls. The authors also reported that they did not find correlation between CBF and DTI parameters in the healthy controls but they observed the correlation in the caudate nucleus in the PD group. However, the study was performed with the images obtained from 1.5T MRI scanner from a limited number of subjects, which were its significant drawback. *The study performed in this thesis used the images obtained from high-resolution 3T MRI scanner on a large cohort of healthy controls and PD patients including Parkinson's disease with normal cognition, Parkinson's disease with mild cognitive impairment and Parkinson's disease dementia (PDD).*

Safai et al. (Safai et al., 2020) performed study on microstructural abnormalities of substantia nigra in Parkinson's disease. Probabilistic atlas of substantia nigra pars compacta was constructed based on neuromelanin sensitive MRI (NMS-MRI) and was utilized to study the microstructural alterations in the SNpc using DTI images obtained from 133 PD and 99 healthy controls. The authors reported that they observed significant increase in AD, MD, and RD values in the pars compacta of PD patients compared to controls while they found no FA differences between the groups. Furthermore, there were no significant correlations demonstrated between DTI measures and clinical scores including UPDRS-III. The authors also reported classification accuracy of 73.4% for the distinction between PD and controls employing a random forest classifier utilizing eight normalized diffusion measures, which included left and right FA, AD, MD, and RD as a feature set. However, the authors did not particularly specify the stages of PD patients used in the study. *This thesis also studies the correlation of DTI measures with global cognitive ability in addition to UPDRS-III on a large group of PD (PD-N, PD-MCI, and PDD). Moreover, we also conducted our study on both substantia nigra and red nucleus, and the segmentations were based on QSM images with the possibility of extension to other deep gray matter nuclei, which was not possible with NMS-MRI because it is specifically suitable for substantia nigra.*

2.4 Chapter Summary

The literature review highlights the successes of previous segmentation methods of the midbrain structures, including substantia nigra and red nucleus, but also areas for improvement, and interesting applications to Parkinson's disease. Currently, atlas-based methods are the most popular method for segmentation of midbrain nuclei. These are performed either in standardized space (warping subjects to normalized space) or in subject space (inverse warping standard regions to subject space) and depend on accurate and robust warping protocols. However, morphometric variability of the midbrain structures limits the accuracy of these segmentation methods. Moreover, most methods proposed in the literature require a substantial number of 'gold standard' reference images for training, which is computationally expensive and not always feasible. *Therefore, work presented in the thesis addresses these issues and focuses on developing the segmentation methods, which are independent of atlas registration, or label-fusion based methods.* The segmentation methods proposed in this thesis also work on both T2*-weighted images and quantitative susceptibility mapping images, making it flexible and potentially multimodal in nature. The literature review also highlighted that most automated segmentation techniques were performed on small number of subjects (including both healthy controls and PD patients) *but this thesis focuses on developing segmentation algorithms and methodology suitable for large cohort of controls and PD patients.*

The review of current literature also revealed a major gap prevalent in this research area, i.e., the disconnection between segmentation techniques and the quantification of brain iron deposition in Parkinson's disease. Almost all the segmentations performed for quantifying brain iron deposition are performed manually. *Therefore, this thesis focuses on developing segmentation algorithms for midbrain structures that correlates well with the gold standard segmentations and use the segmentation results to calculate the susceptibility values and explore their relationship between groups (healthy controls and PD), cognitive function, and motor function in PD.*

The review also provided the overview that different MRI measurements including QSM, R2*, DTI parameters, ASL, and volume reveal distinct tissue characteristics and may indicate different ongoing pathological processes in the midbrain of Parkinson's disease patients. *Therefore, this thesis performs combined assessment of QSM, R2*, FA, AD, MD, RD, CBF, and volume to investigate group differences, and correlation with*

the cognitive and motor impairments. The combined MR parameters were also employed to separate healthy controls and PD. However, this thesis places a greater emphasis on segmentation algorithms and applications to Parkinson's disease than classification algorithms.

3 MATERIALS AND METHODS

This chapter of the thesis presents materials required to conduct the research and methods implemented to achieve the goals. MRI acquisition processes for different databases used for the experimental analysis, diagnostic criteria and assessment of our own database, and MRI preprocessing are discussed. This chapter also briefly describes quantitative assessment performed to gauge the segmentation accuracy and correlation coefficients used for the evaluation of the results.

3.1 Materials

The research work in this thesis uses two distinct MRI databases (i) online database which is freely available and described in detail by Forstmann et al. (Forstmann et al., 2014a) and (ii) the Parkinson's disease cohort at the New Zealand Brain Research Institute (NZBRI).

3.1.1 Online Database

This research uses a publicly available MRI data repository (Forstmann et al., 2014b) consisting of scans from healthy participants; young, middle-aged and elderly. The data were obtained using a 7T Siemens Magnetom MRI scanner (Siemens Healthcare, Erlangen, Germany) using a 24-channel head array Nova coil (NOVA Medical Inc., Wilmington MA). The multi-echo 3D FLASH consisted of 128 slices with repetition time (TR) = 41 ms, three echo times (TE): 11.22/20.39/29.57 ms, bandwidth = 160 Hz/Px, flip angle = 14 deg, voxel size = $(0.5\ \text{mm})^3$ and was tilted at -23 deg.

For the experiments in Chapter 4, we use the online database. The sample scan from online database is displayed in Figure 1.2.

3.1.2 New Zealand Brain Research Institute Database

A convenience sample of 127 participants meeting the UK Parkinson's Disease Society's criteria for idiopathic PD (Hughes, Daniel, Kilford, & Lees, 1992) was recruited from volunteers at the Movement Disorders Clinic at the New Zealand Brain Research Institute, Christchurch, New Zealand. Exclusion criteria included atypical Parkinsonian disorders; prior learning disability; previous history of other neurological conditions including moderate-severe head injury, stroke, vascular dementia; and major psychiatric or medical illness in the previous 6 months. 47 healthy controls matched to the PD group by mean age and sex ratio, were also recruited.

Participants completed neuropsychological testing, clinical assessment, and an MRI scanning session. All participants gave written informed consent, with additional consent from a significant other when appropriate. The study was approved by the regional Ethics Committee of the New Zealand Ministry of Health (No. URB/09/08/037).

For the experimental work in Chapters 5 and 6 we use MRI images from the NZBRI database. Sample scans from NZBRI database are displayed in Figure 1.1, Figure 1.3, Figure 1.5, and Figure 1.6.

3.1.2.1 Diagnostic Criteria and Assessment

Motor function was assessed using the Movement Disorders Society Unified Parkinson's Disease Rating Scale part 3 (Goetz et al., 2008). Comprehensive neuropsychological assessment fulfilling the Movement Disorders Society Task Force Level II criteria was used to diagnose PD-MCI (Litvan et al., 2012; Wood et al., 2016). Five cognitive domains were examined (executive function; attention, working memory and processing speed; learning and memory; visuospatial/visuoperceptual function; and language). Within each cognitive domain, standardized scores from the constituent neuropsychological tests were averaged to provide individual cognitive domain scores; global cognitive performance for each participant was expressed as an aggregate z score obtained by averaging across the five domain scores. PD-MCI cases had unimpaired functional activities of daily living, as verified by interview with a significant other, and scored 1.5 standard deviation (SD) or more below normative data in at least one of the

five cognitive domains (Wood et al., 2016). Dementia was defined using MDS criteria as significant cognitive impairments (2 SD below normative data) in at least two of five cognitive domains, plus evidence of significant impairment in everyday functional activities, not attributed to motor impairments (Emre et al., 2007). Participants also completed the Montreal Cognitive Assessment. All assessments and scans were performed with no disruption to participants' usual medication regimen. PD participants were classified as either cognitively normal (PD-N, n = 46), with mild cognitive impairment (PD-MCI; n = 66), or with dementia (PDD; n = 15).

3.1.2.2 MRI Acquisition

All 174 participants (47 healthy controls and 127 PD patients) were scanned on a 3T General Electric HDxt scanner (GE Healthcare, Waukesha, USA) with an eight channel head coil. To quantify magnetic susceptibility in the brain, a 3D spoiled gradient recalled echo acquisition with 8 echoes was used to obtain real and imaginary pairs with the following parameters: TE = 3.5, 7.3, 11.1, 14.9, 18.7, 22.5, 26.3, 30.2 ms (3.8 ms intervals), TR = 42.8 ms, flip angle = 20 deg, acquisition matrix = 512×512×60, field of view (FOV) = 240 mm, slice thickness = 2 mm, voxel size = 0.47×0.47×2 mm^3. In addition, we acquired a conventional T1-weighted 3D SPGR acquisition (TE/TR = 2.8/6.6 ms, inversion time (TI) = 400 ms, flip angle = 15 deg, acquisition matrix = 256×256×170, FOV = 250 mm, slice thickness = 1 mm, voxel size = 0.98×0.98×1.0 mm^3), a T2-weighted, and T2 weighted fluid-attenuated inversion recovery images to facilitate a clinical read. All MRI scans were screened by a consultant neuroradiologist (RJK) to exclude significant non-PD cerebral pathology.

3.1.2.2.1 Diffusion Tensor Imaging

A 2D diffusion-weighted, spin echo, echo planar imaging sequence was used to measure microstructural integrity, with diffusion weighting in 28 uniformly distributed directions (b = 1000 s/mm^2) and 4 acquisitions without diffusion weighting (b = 0 s/mm^2): TE/TR = 86.4/13000 ms, flip angle = 90 deg, acquisition matrix = 128×128×48, reconstruction matrix = 256×256×48, FOV = 240 mm, slice thickness = 3 mm, reconstructed voxel size = 1.07×1.07×3 mm^3, NEX = 1, ungated.

3.1.2.2.2 Arterial Spin Labelling

A stack of spiral, fast spin echo acquired images were prepared with pseudo-continuous arterial spin labelling and background suppression to measure whole brain

perfusion (Dai et al., 2008) quantitatively: TR = 6 s, echo spacing = 9.2 ms, post-labelling delay = 1.525 s, labelling duration = 1.5 s, eight interleaved spiral arms with 512 samples at 62.5 kHz bandwidth and 30 phase encoded 5mm thick slices, NEX = 5, units: ml/100 g/min). Participants were asked to close their eyes during the ASL acquisition.

3.1.2.3 MRI Preprocessing

3.1.2.3.1 QSM Reconstruction

The morphology enabled dipole inversion (MEDI) algorithm was used to generate quantitative susceptibility maps from the real/imaginary pairs (T. Liu, Wisnieff, et al., 2013; T. Liu, Xu, Spincemaille, Avestimehr, & Wang, 2012; Z. Liu, Spincemaille, Yao, Zhang, & Wang, 2018). The total field was estimated using a nonlinear field map estimation followed by graph-cut-based phase unwrapping (Dong et al., 2015). We used the Laplacian-boundary value technique for background field removal (Zhou, Liu, Spincemaille, & Wang, 2014) and morphology enabled dipole inversion with automatic uniform ventricular cerebrospinal fluid zero referencing (MEDI+0) to produce the quantitative susceptibility maps. A brain mask was extracted from the averaged magnitude image using FSL's (v6.0.3) brain extraction tool (BET) (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012).

CAT12 (r1278, http://www.neuro.uni-jena.de/cat/), a toolbox of SPM12 (v7219, http://www.fil.ion.ucl.ac.uk/spm/), running in MATLAB 9.3.0 (R2017b), was used to process T1-weighted structural images. Each individual's T1-weighted structural image was first re-oriented (using SPM12's coregistration routine) to be in rough alignment with Montreal Neurological Institute (MNI) space; the average magnitude image was then coregistered to the T1-weighted structural image, along with each individual QSM map, to ensure a relatively consistent orientation in subject space.

T1-weighted images were bias corrected, spatially normalized via Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) (using the MNI-registered template provided within CAT12), modulated to compensate for the effect of spatial normalization, and classified into gray matter, white matter, and cerebrospinal fluid (CSF), all within the same generative model (Ashburner & Friston, 2005). QSM images were warped into MNI space using the T1-derived deformation fields and averaged to create a study-specific, average QSM image.

We used QSM images for the experimental analysis in Chapters 5 and 6.

3.1.2.3.2 DTI Preprocessing

Diffusion-weighted images were pre-processed using FSL (v6.0.3, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). This included motion- and eddy-current distortion-correction; rotation of the diffusion direction encoding ('b') matrix accordingly; motion quantification via root mean square deviation between each pair of realigned diffusion images and averaging over all pairs to create a single, 'relative' motion metric; brain extraction; and fitting a diffusion tensor to produce fractional anisotropy, mean diffusivity, axial diffusivity (the principal diffusion eigenvalue), and radial diffusivity (the mean of the second and third eigenvalues) images. The raw T2 image with no diffusion weighting (the b = 0 s/mm^2 image) was then non-linearly coregistered and warped into T1-weighted structural space (using CAT12 non-linear coregistration); these parameters were used to warp FA, MD, AD, and RD images into T1 space, as well.

3.1.2.3.3 ASL Preprocessing

Quantified cerebral blood perfusion images were co-registered to the structural T1-weighted image (using SPM12 routines).

3.2 Target Region Selection

In order to begin the segmentation, the region containing substantia nigra and red nucleus (target midbrain nuclei) must be selected from input brain MRI (T2*-weighted image and QSM image). The target region was obtained manually in Chapter 4 and through substantia nigra seeding in Chapter 5 by creating a bounding box to initiate SN and RN segmentation. The details of substantia nigra seeding will be described later in Chapter 5.

3.3 Subjective Assessment

Qualitative assessment of segmentations in Chapters 4 and 5 were separately performed by radiologists' (RJK, AL, and MH) having years of experience in neuroimaging. Subjective evaluation was performed blindly, meaning the segmentation result of the methods proposed in the thesis and the various other methods used for comparison were mixed up and unknown to the radiologists'. Subjective assessment

was performed using the score range from 1-5 (very poor, poor, fair, good, very good) for the obtained segmentation results.

Some of the major point radiologists' took into account while scoring the segmentations are listed as follows:

(i) How well the segmentation contour was obtained within the boundary of the SN and RN. A higher subjective score was given to the segmentation results that could tightly fit the midbrain nuclei whereas a significantly lower score was given to the segmentation outputs that extend beyond the boundary of SN and RN.

(ii) Over-segmentation present in the outputs including holes in the midbrain nuclei were also given a lower score.

(iii) The smoothness of the segmentation contour was also considered.

(iv) The separation of SN and RN was also emphasized; a lower score was given to the segmentations where SN and RN were intact together.

3.4 Manual Segmentations

Manual segmentation of substantia nigra and red nucleus in Chapters 4 and 5 was performed by experienced radiologist using the FSLeyes (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLeyes) software for each subject. The radiologist segmented SN and RN on both left and right sides. The manual segmentation was considered as the gold standard for further analysis.

3.5 Assessment of Segmentation Accuracy

The segmentation accuracy is assessed using Dice coefficient (Dice, 1945) which is a commonly used similarity metric to measure the fraction of spatial overlap between the two segmentations; specifically, manual segmentations performed by the radiologists were compared with the results generated by our segmentation methods (Chapters 4 and 5). Mathematically, Dice coefficient is given by,

$$Dice = 2\frac{|G \cap A|}{|G| + |A|} \qquad \textbf{3.1}$$

Where *G* and A denote gold standard segmentations and automated segmentations respectively and $\cap$ is the intersection operator.

The value of Dice score ranges between 0 and 1; 0 means that there is no overlap between the two segmentation results ($G \cap A = \phi$) while 1 means that the two segmentation results completely overlap with each other (G = A).

3.6 Correlation Coefficients

Correlation of QSM values extracted from the proposed segmentations and gold standard segmentations of substantia nigra and red nucleus is assessed through Pearson correlation coefficient and intraclass correlation coefficient. They are discussed in brief below.

3.6.1 Pearson Correlation Coefficient

Pearson correlation coefficient of QSM values from gold standard and proposed segmentation method is defined as the ratio of the covariance of QSM values from both the methods to the product of their individual standard deviations. Mathematically, Pearson correlation coefficient (*r*) is given by,

$$r = \frac{\sigma_{MN}}{\sigma_M \sigma_N} \qquad \textbf{3.2}$$

Where σ_{MN} is the covariance of QSM values from gold standard segmentations and proposed method, σ_M is the standard deviation of gold standard segmentations, and σ_N is the standard deviation of proposed segmentations.

The values of Pearson correlation coefficient ranges from -1 to +1; $r = 1$ signifies perfect positive linear association, $r = -1$ indicates perfect negative linear association, and $r = 0$ signifies a weak linear association between the QSM values obtained from both the segmentations.

3.6.2 Intraclass Correlation Coefficient

Intraclass correlation coefficient provides the reliability measurements for the QSM values obtained from proposed automated segmentations and gold standard segmentations. ICC reflects both degree of agreement and correlation between the

methods. ICC is also widely used in interrater reliability analyses. There are several versions of ICC (McGraw & Wong, 1996) but we will discuss type (A,1) and (C,1) in particular, as we use them in our research in Chapters 4 and 5.

3.6.2.1 ICC(A,1)

ICC(A,1) measures the degree of absolute agreement among the measurements. Mathematically,

$$ICC(A,1)=\frac{MS_R-MS_E}{MS_R+(k-1)MS_E+\frac{k}{n}\left(MS_c-MS_E\right)} \quad \textbf{3.3}$$

3.6.2.2 ICC(C,1)

ICC(C,1) measures the degree of consistency among the measurements. Mathematically,

$$ICC(C,1)=\frac{MS_R-MS_E}{MS_R+(k-1)MS_E} \quad \textbf{3.4}$$

Where MS_R is the mean square for rows, MS_E is the mean square error, MS_c is the mean square for columns, n is the number of subjects and k is the number of measurements.

ICC values less than 0.5 indicates poor reliability, 0.5 – 0.75 indicate moderate reliability, 0.75 – 0.9 indicate good reliability, and values larger than 0.9 indicate excellent reliability (Koo & Li, 2016).

3.7 Chapter Summary

This chapter presented an overview of the MRI databases used to perform the research carried out in this thesis. QSM reconstruction as well as the subjective assessment of the segmentation was described. The quantitative assessment technique followed to estimate the segmentation accuracy in Chapters 4 and 5 was discussed, and the correlation measures adopted in Chapter 5 briefly described. The next chapter describes two proposed methodologies for the segmentation of substantia nigra using T2*-weighted images acquired from an online database.

4 SEGMENTATION OF SUBSTANTIA NIGRA IN T2*-WEIGHTED IMAGES

Conventional magnetic resonance imaging at standard field strength fails to visualize the structure of substantia nigra because of the insufficient signal-to-noise ratio and spatial resolution. But, advanced MR imaging technique at 3T and 7T have improved signal-to-noise ratio to increase the spatial resolution and therefore smaller brainstem nuclei like SN can be visualized. Additionally, the usage of gradient-recalled-echo sequences helps in characterizing SN which contains relatively high iron content (Hallgren & Sourander, 1958; Sofic, Paulus, Jellinger, Riederer, & Youdim, 1991). SN is not clearly visible in T1-weighted images while it is readily visible in T2-weighted images and T2*-weighted images because of the presence of iron.

Accurate segmentation of substantia nigra located in midbrain is particularly important to both clinicians and researchers because iron deposition in SN is found to be highly correlated with Parkinson's disease. However, there are limited techniques available for the segmentation (Garzón et al., 2018; Guo et al., 2018; Haegelen et al., 2013; Kim et al., 2014; Visser et al., 2016; Yiming Xiao et al., 2012; Y. Xiao et al., 2014) of the midbrain nuclei. In view of the limited literature, this chapter presents two improved segmentation algorithms for the effective segmentation of substantia nigra in T2*-weighted images: (i) Segmentation method based on weighted thresholding (ii) Segmentation based on level set method and wavelet transform. The segmentation based

on weighted thresholding uses different samples of the image to estimate the threshold value suitable for efficient segmentation of substantia nigra. Similarly, segmentation based on the combination of level set method and wavelet transform uses level set method to obtain the segmentation contour of the midbrain nuclei while wavelet transform is employed for improving and addressing the issues in segmentation outputs generated by the level set methods. The experiments in this chapter were performed using the online database (Forstmann et al., 2014b) utilizing T2*-weighted images from 2 young, 2 middle-aged and 2 elderly participants. This chapter also presents the experimental results generated by both segmentation methods and validation performed by clinical experts' through qualitative and quantitative assessment. This thesis also presents the comparative analysis to the various state-of-the-art algorithms to show the effectiveness of the proposed methodologies.

4.1 Segmentation based on Weighted Thresholding

The flowchart of the proposed method is shown in Figure 4.1. To initiate the segmentation, the region containing substantia nigra (target nuclei) was selected manually from the input brain MRI by obtaining a bounding box in the target region. Then, *k*-means algorithm was applied on different samples of target region including cross-diagonal pixels (Zortea, Flores, & Scharcanski, 2017) and the pixels outside the cross-diagonals (Figure 4.2 (b)) to compute the threshold value along each of these samples. The threshold estimated along these samples were weighted to calculate the final threshold value for the accurate segmentation of substantia nigra. The details of the proposed methodology are discussed next.

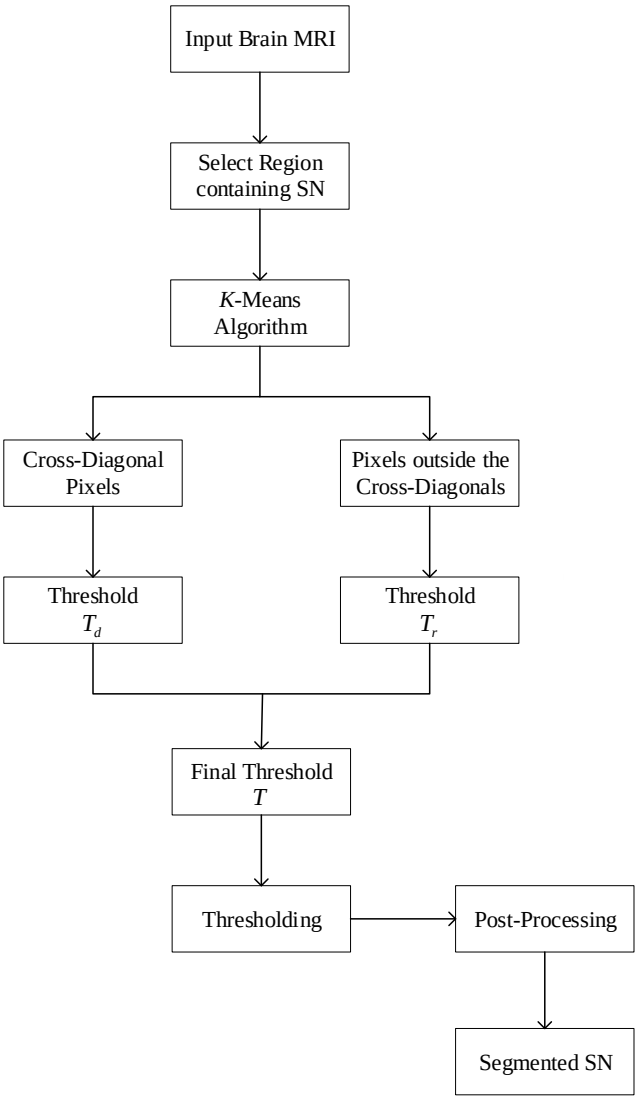


Figure 4.1: Flowchart of the proposed method.

4.1.1 K-Means Algorithm

K-means (Jain, 2010; Laszlo & Mukherjee, 2006; J. Liu & Guo, 2015; Ng, Ong, Foong, Goh, & Nowinski, 2006) algorithm is a popular unsupervised method which is widely used because of its simplicity. It partitions the target region intensities into *k* clusters and aims to minimize the sum of squared distance between the data points and the cluster center. The main steps of the *k*-means algorithm to partition the target region intensities y_i, $i=1,2,3....m$ into *k* clusters are as follows.

1. Initialize cluster number *k* and cluster center C_j.

2. The data point y_i is assigned to the group that has the closest cluster center.

3. Compute new cluster centers after assigning all the data points to any of the clusters.
4. Steps 2 and 3 are repeated until cluster membership stabilizes.

The goal of *k*-means algorithm is to minimize an objective function given by,

$$J = \sum_{j=1}^{k} \sum_{i=1}^{m} \left\| y_i - C_j \right\|^2 \qquad \textbf{4.1}$$

where $\left\| y_i - C_j \right\|$ is the Eucledian distance between a data point y_i and centroid C_j.

K-means algorithm is an efficient method for data clustering and offers a low computational complexity. It is also known as hard classification algorithm because it forces every pixel to belong to one and only one class in each iteration unlike in soft classification methods (Kalti, 2012; Shan, Sandham, Granat, & Sterr, 2005). Therefore *k*-means algorithm is also referred to as a simpler non-probabilistic alternative to Gaussian mixture.

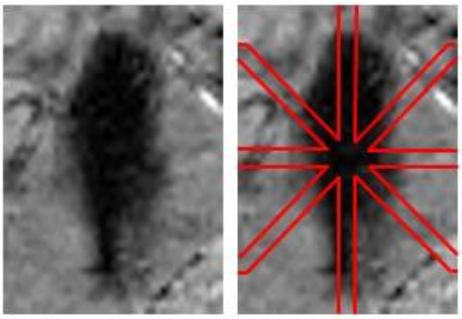


Figure 4.2: Image representation for (a) all pixels (b) only cross-diagonal pixels.

K-means algorithm was applied on pixel values using $k=2$ as the image predominantly consists of two regions as shown in Figure 4.2 (a). In our work, we used *k*-means algorithm along only cross-diagonal pixels and the pixels outside the cross-diagonals. The image representation of the proposed approach is shown in Figure 4.2 (b). The red lines represents the cross-diagonal pixels and were selected in such a way that it passed through the center of the image. But, the cross-diagonal pixels did not include entire SN, and hence could not segment them accurately. Therefore, the pixels outside the cross-diagonals were also considered for accurate segmentation. *K*-means was

performed for both the samples of the image. The algorithm resulted in centroids represented by C_1 and C_2 for each region. The threshold (Huang & Chau, 2008; D. J. Liu & Yu, 2009; Zortea et al., 2017) T' is calculated as,

$$T' = \frac{C_1 + C_2}{2} \quad \textbf{4.2}$$

Where $T' = T_d$ represents cross-diagonal pixels and $T' = T_r$ represents pixels outside the cross-diagonals.

The final threshold value T using our proposed method is calculated as,

$$T = \alpha T_r + (1 - \alpha) T_d \quad \textbf{4.3}$$

Where α is an empirically determined constant in the range [0, 1] which balances the two thresholds.

4.1.2 Thresholding

Thresholding is a simple, fast, and computationally efficient method for image segmentation because it differentiates the gray level of pixels belonging to the object from the gray level of the pixels belonging to the background. Thresholding can be single or multilevel thresholding. In our work, we use single threshold to separate the image into different classes.

Substantia nigra varies in shape and size as shown in Figure 4.3 (a). Initially, we estimated threshold using all the pixels in the image employing *k*-means algorithm with $k = 2$ which resulted in two centroids. The threshold for all image pixels was computed as the average of the centroids. However, the threshold thus obtained segments SN as shown in Figure 4.3 (b), with the boundary containing some of the irrelevant background pixels. Therefore, we estimated threshold value employing our proposed method utilizing equation (4.3) to categorize the image regions into their respective classes for the effective segmentation. The result of thresholding operation is always a binary image as shown in Figure 4.4 (b).

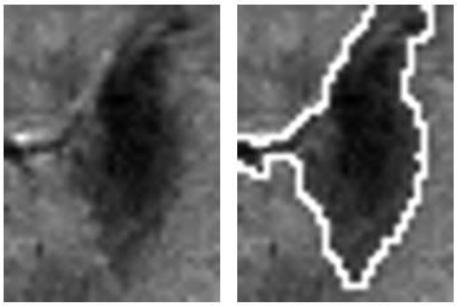


Figure 4.3: (a) Input image (b) Under-segmented image. *K*-means threshold for all image pixels = 68.

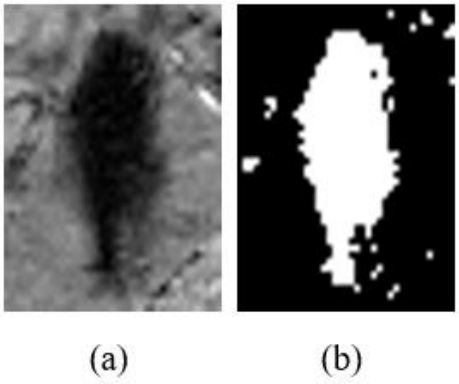


Figure 4.4: (a) Input image (b) Binary image. Threshold estimated using our proposed method, T = 0.8 · 45.6 + (1 – 0.8) · 38.9 = 44.3 (α = 0.8).

4.1.3 Post-Processing

The binary image obtained after the thresholding stage contains many disjoint pixels as shown in Figure 4.4 (b) i.e., some of the background pixels are misclassified as the foreground objects. Therefore, the goal of the post-processing stage was to remove all the disjoint objects and select the single largest connected region that is likely to be SN, which we are trying to segment. The single connected region consisted of holes, which needs to be filled using morphological reconstruction. Moreover, the boundary of the single connected region was not smooth; hence, it was smoothened by a Gaussian filter. We used two-dimensional Gaussian low pass filter of size 9×9 for this purpose. Finally, we detect the border of the connected region and overlay the input image with the detected border to obtain the desired segmentation result.

4.2 Segmentation based on Level Set Method and Wavelet Transform

Level set method originally proposed by Osher and Sethian (Osher & Sethian, 1988) is a popular technique for capturing the evolution of dynamic interfaces. It represents the interfaces or contours as the zero level set of a higher dimensional function, usually known as a level set function. Level set method is capable of handling the changes in topology of the interface and shapes in a natural way. On the other hand, wavelet transform is a popular method for smoothing the images (Mallat, 1999). Therefore, this section of the thesis describes the proposed segmentation method based on level set and wavelet transform in detail.

4.2.1 Level Set Method

Chan and Vese (CV) (Chan & Vese, 2001) proposed a level set method based on the Mumford-Shah model (Mumford & Shah, 1989) for minimizing the energy given by,

$$F^{CV}(C,c_1,c_2)=\lambda_1\int_{\Omega_1}\left|I(x)-c_1\right|^2dx+\lambda_2\int_{\Omega_2}\left|I(x)-c_2\right|^2dx+v\left|C\right| \qquad \textbf{4.4}$$

Where $v\geq 0$, $\lambda_1,\lambda_2>0$, Ω_1 and Ω_2 are the regions outside contour *C* and inside contour *C* respectively, c_1 and c_2 approximate the image intensity in Ω_1 and Ω_2. However, the CV model failed to segment images with intensity inhomogeneities, common in medical images (C. Li, Kao, Gore, & Ding, 2008).

Hence, Li et al. (C. Li et al., 2008; C. M. Li, Kao, Gore, & Ding, 2007) addressed this issue by embedding local image information, which helps segmenting images with intensity inhomogeneities. This model was called the region-scalable fitting energy model. The energy functional for RSF model is given by,

$$\begin{aligned}F(\phi,f_1,f_2)=&\sum_{i=1}^{2}\lambda_i\int\left(\int K_\sigma(x-y)\left|I(y)-f_i(x)\right|^2M_i^{\varepsilon}(\phi(y))dy\right)dx\\&+v\int\left|\nabla H_\varepsilon(\phi(x))\right|dx+\mu\int\frac{1}{2}\left(\left|\nabla\phi(x)\right|-1\right)^2dx\end{aligned} \qquad \textbf{4.5}$$

Where $f_1(x)$ and $f_2(x)$ approximate the local intensities in the regions outside and inside the contour, respectively. K_σ is a Gaussian kernel with a scale parameter $\sigma>0$. H_ε is a smooth Heaviside function, $M_1^\varepsilon(\phi)=H_\varepsilon(\phi)$ and $M_2^\varepsilon(\phi)=1-H_\varepsilon(\phi)$, μ is a positive constant.

$$H_\varepsilon(x)=\frac{1}{2}\left[1+\frac{2}{\pi}\arctan\left(\frac{x}{\varepsilon}\right)\right] \quad \textbf{4.6}$$

The minimization of energy functional $F(\phi, f_1, f_2)$ was performed with respect to ϕ (keeping f_1 and f_2 fixed), using the standard gradient descent method by solving the gradient flow equation as:

$$\frac{\partial\phi}{\partial t}=-\delta_\varepsilon(\phi)(\lambda_1 e_1-\lambda_2 e_2)+v\delta_\varepsilon(\phi)div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)+\mu\left(\nabla^2\phi-div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right) \quad \textbf{4.7}$$

Where δ_ε is the smoothed Dirac delta function.

$$\delta_\varepsilon(x)=\frac{1}{\pi}\frac{\varepsilon}{\varepsilon^2+x^2} \quad \textbf{4.8}$$

Also,

$$\begin{cases} e_1(x)=\int K_\sigma(y-x)|I(x)-f_1(y)|^2 dy \\ e_2(x)=\int K_\sigma(y-x)|I(x)-f_2(y)|^2 dy \end{cases} \quad \textbf{4.9}$$

And,

$$\begin{cases} f_1(x)=\dfrac{K_\sigma*[H_\varepsilon(\phi)I]}{K_\sigma*H_\varepsilon(\phi)} \\ f_2(x)=\dfrac{K_\sigma*[(1-H_\varepsilon(\phi))I]}{K_\sigma*(1-H_\varepsilon(\phi))} \end{cases} \quad \textbf{4.10}$$

Also,

$$\lambda_1 e_1 - \lambda_2 e_2 = (\lambda_1 - \lambda_2) I^2 [K_\sigma * 1] - 2I[K_\sigma * (\lambda_1 f_1 - \lambda_2 f_2)] + K_\sigma * (\lambda_1 f_1^2 - \lambda_2 f_2^2) \quad \textbf{4.11}$$

Where $*$ is the convolution operation.

In the proposed level set method, we aim to solve level set equation given by (4.7). The first term in equation (4.7) drives the active contour towards the boundaries of the object while the second term maintains the regularity of the contour. The third term in equation (4.7) is called a level set regularization term. This term helps to avoid the time-consuming re-initialization process by maintaining the regularity of the level set function; this allows for accurate computation and stable evolution in level set methods.

4.2.2 Wavelet Transform

In this work, we use wavelet transform for smoothing the images. Wavelet based smoothing is performed by using a threshold to remove high frequency subband coefficients. Thresholding can be performed either by hard or soft thresholding method. In the proposed work, dual-tree complex wavelet transform is used.

4.2.2.1 Dual-Tree Complex Wavelet Transform

Dual-tree complex wavelet transform (Kingsbury, 1998, 2001) proposed by Kingsbury is an improvement to the conventional discrete wavelet transform (DWT). DWT is shift variant which leads to the significant change of wavelet coefficients at the output even with a small shift of the input signal. Additionally, DWT produces only three directionally selective subbands ($0^\circ, 45^\circ, 90^\circ$). Therefore, to overcome these limitations, DT-CWT was proposed which is shift invariant and improves directional resolution by producing six directionally selective subbands ($\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$). Moreover, DT-CWT is robust to noise and performs perfect reconstruction.

The one-dimensional (1-D) DT-CWT (Celik & Tjahjadi, 2010) is shown in Figure 4.5 which uses two DWTs. The first DWT gives the real part of the transform, while the second DWT gives the imaginary part. A separate filter pair is used with $h_0(n)$ and $h_1(n)$ for the real part and $g_0(n)$ and $g_1(n)$ for the imaginary part. The 1-D DT-CWT decomposes an input signal $f(x)$ in terms of a complex shifted and dilated mother wavelet $\psi(x)$ and scaling function $\varphi(x)$. Mathematically,

$$f(x)=\sum_{l\in Z} B_{j_0,l}\varphi_{j_0,l}(x)+\sum_{j\geq j_0}\sum_{l\in Z} w_{j,l}\psi_{j,l}(x) \qquad \textbf{4.12}$$

Where Z is the set of natural numbers, j_0 represents the number of decomposition level, j and l are the index of shifts and dilations, respectively. $B_{j_0,l}$ is the scaling coefficient and $w_{j,l}$ represents the complex wavelet coefficient. $\varphi_{j_0,l}(x)=\varphi^r_{j_0,l}(\mathrm{x})+i\varphi^i_{j_0,l}(\mathrm{x})$ represents the scaling function and $\psi_{j,l}(x)=\psi^r_{j,l}(\mathrm{x})+i\psi^i_{j,l}(\mathrm{x})$ is the complex wavelet function where the superscripts r and i represent the real and imaginary parts, respectively.

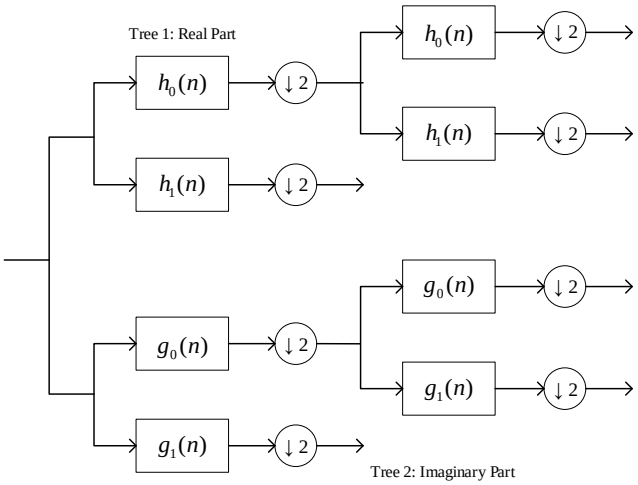


Figure 4.5: 1-D dual-tree complex wavelet transform.

Similarly, the two-dimensional (2-D) DT-CWT decomposes an image $f(x,y)$ through a sequence of dilations and translations of a complex scaling function and six complex wavelet functions. Mathematically,

$$f(x,y)=\sum_{l\in Z^2} B_{j_0,l}\varphi_{j_0,l}(x,y)+\sum_{\theta\in\beta}\sum_{j\geq j_0}\sum_{l\in Z^2} w^{\theta}_{j,l}\psi^{\theta}_{j,l}(x,y) \qquad \textbf{4.13}$$

Where $\theta\in\beta=\left\{\pm 15^{\circ},\pm 45^{\circ},\pm 75^{\circ}\right\}$ gives the directionality of the complex wavelet function. Therefore, the decomposition of a 2-D image $f(x,y)$ by DT-CWT gives a complex-valued low pass subband and six complex-valued high pass subbands at every level of decomposition, where each high pass subband represents one unique direction θ.

4.2.3 Implementation

This section describes the implementation detail of the segmentation based on level set method and wavelet transform.

In the proposed level set method, all the partial derivatives $\frac{\partial \phi}{\partial x}$ and $\frac{\partial \phi}{\partial y}$ in equation (4.7) were discretized as central finite differences while the temporal derivative was discretized as a forward difference. Iteration scheme was subsequently obtained by discretizing the partial differential equation (PDE) in equation (4.7). The binary step function was used for the initialization of the level set function as it facilitates easier emergence of new contours and faster curve evolution than the initialization with a signed distance map. The binary step function takes negative constant value $-c_0$ inside the region and positive constant value c_0 outside the region. In our experiments in this thesis, we used $c_0 = 2$. Moreover, the functions f_1 and f_2 were updated at every time step as per equation (4.10) before the update of the level set function.

Gaussian kernel K_σ with a scale parameter σ was used as a convolution kernel. The convolution kernel was constructed as a $w \times w$ mask, such that $w \geq 4 * \sigma + 1$. In our work, we used $\sigma = 10$, i.e., the size of mask was 41×41. A smaller σ may be used, but it would require a large number of iterations for accurate computation.

In equation (4.10), there were a total of four convolutions across the numerator and denominator. The two convolutions, $K_\sigma * I$ and $K_\sigma * 1$, required to compute f_2 were determined only once before the iterations as the evolution of level set function did not depend on them. But, the other two convolutions, $K_\sigma * [H_\varepsilon(\phi) I]$ and $K_\sigma * H_\varepsilon(\phi)$, required for computing the functions f_1 and f_2 needed to be computed at every iteration for evolving the level set function. Similarly, $(\lambda_1 e_1 - \lambda_2 e_2)$ can be written as a combination of three convolutions as shown in equation (4.11). However, one convolution was not dependent on evolving the level set function and could be computed only once before the iterations, like in the previous case. Therefore, we needed to compute four convolutions at each iteration for evolving the level set function.

However, the level set method produced jagged segmentation outputs, not the desired smooth segmentation outputs. Moreover, it also over-segmented and under-segmented

the midbrain nuclei. Therefore, DT-CWT was used to address these issues and employed while evolving the level set function. DT-CWT consisted of two wavelets in each direction and each wavelet was oriented in six distinct directions. Therefore, one wavelet could be considered as the real part while the other could be interpreted as the imaginary part of the complex-valued wavelet. In our experiments, we took forward DT-CWT over 3 levels of decomposition, with different filter sets along the rows and columns. Oriented wavelets were obtained by performing sum and difference of wavelets along the x and y directions. Moreover, sum and difference operation was normalized by $1/\sqrt{2}$. Additionally, we thresholded the high frequency wavelet coefficients through all scales and sub-bands using a heuristic approach. Finally, we took the inverse DT-CWT to resolve the issues generated from level set method.

4.3 Qualitative Assessment of Segmentations

Subjective evaluation was performed separately and blindly by two experienced radiologists' (RJK and AL) for the assessment of the segmentations produced by the proposed methods and the various other methods used for comparison. A score ranging from 1-5 was used for the subjective assessment of the obtained segmentation results.

4.4 Quantitative assessment: Expert ('Gold standard') Manual Segmentations

Manual segmentations of substantia nigra was performed by an experienced radiologist (AL) for all 6 subjects. The manual segmentation was considered as the gold standard for performing the quantitative analysis.

Dice score given in equation (3.1) was used to estimate the degree of overlap between gold standard segmentations and proposed segmentations.

4.5 Comparison with Different Methods

We compared the proposed segmentation method based on weighted thresholding with threshold estimated from all the pixels in the image. *K*-means algorithm ($k=2$) was applied resulting in centroids C_1 and C_2 as discussed in Section 4.1.2. The darker region corresponds to the substantia nigra as shown in Figure 4.2. Therefore, we assumed the centroid of the darker cluster as the threshold for fair comparison.

Segmentation based on level set method and wavelet transform was compared with the popular level set methods to show the effectiveness of our proposed algorithm in all the 6 subjects because it is frequently used method for the segmentation of smaller structures in medical imaging. We used level set method based on the minimization of region scalable fitting energy for segmentation of images with intensity inhomogeneities proposed by Li et al. (C. Li et al., 2008) for comparison. Similarly, the level set method proposed by Li et al. (C. Li et al., 2011) which is based on the local intensity clustering property and is capable of handling images with intensity inhomogeneities was also used for comparison.

4.6 Experimental Results and Discussion

At first, we will describe in detail, the results of segmentation based on weighted thresholding approach including its advantages and limitations followed by the method based on level set and dual-tree complex wavelet transform.

4.6.1 Algorithm Performance (Weighted Thresholding)

The segmentation results of the proposed weighted thresholding method, comparison with threshold estimated from all the pixels in the image and the darker cluster, and manual segmentations of the SN are shown in Figure 4.6. For all the experiments, we use $\alpha = 0.80$. The value of α was defined after evaluation with several variations. The segmentation result of the right and left side of the SN clearly shows that the proposed segmentation method produced good results and the labeled mask of SN were closer to the manual segmentation. However, the threshold calculated using all the pixels of the image and the darker cluster could not segment SN effectively and efficiently.

Our proposed segmentation method obtained a higher subjective score from both radiologists' in comparison to the other methods in the blind assessment and it validates our methodology. The subjective score from the radiologists' for the sample scans presented in Figure 4.6 are shown in Table 4.1 (RJK) and Table 4.2 (AL). The radiologists' stated that the proposed methodology enclosed better SN in comparison to the other two methods. The inter-rater agreement between the subjective scores (ICC) is also presented in Table 4.3.

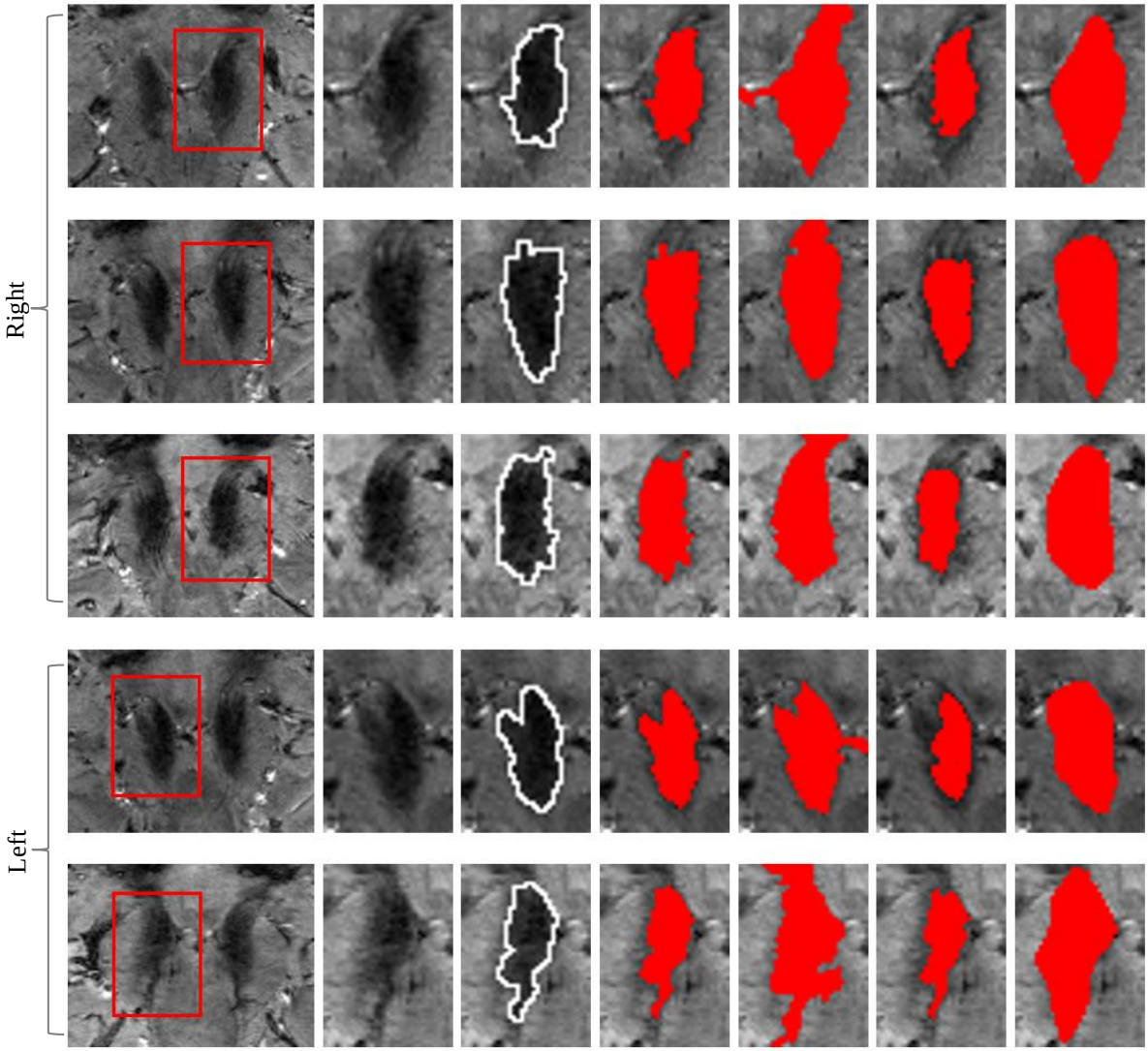


Figure 4.6: Results of the proposed segmentation method and comparison with different methods. Column 1: Bounding box selected manually. The box is represented in red. Column 2: Area within the bounding box to initiate the segmentation. Column 3: Segmentation result of our proposed method. The result of segmentation is shown as a white contour. Column 4: labeled mask of the proposed segmentation method is shown in red (SN). Column 5: Result of the threshold estimated from all the pixels in the image. Column 6: Result of the threshold value of the darker cluster. Column 7: Manually labeled structures of SN.

Table 4.1: Subjective Evaluation Scores (Radiologist 1: RJK)

	Scan Number	Proposed Method	All Image Pixels	Darker Cluster
Right	Scan 1	3	1	2
	Scan 2	4	1	2
	Scan 3	3	2	1
Left	Scan 4	3	3	2
	Scan 5	3	1	2

Score 1: very poor; Score 2: poor; Score 3: fair; Score 4: good; Score 5: very good

Table 4.2: Subjective Evaluation Scores (Radiologist 2: AL)

	Scan Number	Proposed Method	All Image Pixels	Darker Cluster
Right	Scan 1	3	2	2
	Scan 2	4	2	3
	Scan 3	4	3	2
Left	Scan 4	3	3	2
	Scan 5	3	2	2

Score 1: very poor; Score 2: poor; Score 3: fair; Score 4: good; Score 5: very good

Table 4.3: Inter-rater agreement between the subjective scores

Method		
Proposed Method	0.57	Moderate
All Image Pixels	0.32	Poor
Darker Cluster	0.09	Poor

The Dice score for the proposed method, all image pixels, and the darker cluster was also calculated for the left and right side of the SN for all the 6 subjects and is shown in Figure 4.7. The Dice score (mean (standard deviation)) of the left SN and right SN between the manual segmentation and the proposed method were 0.71 (0.08) and 0.69

(0.06) respectively; Dice for all image pixels were 0.66 (0.04) and 0.56 (0.07) while the Dice for the darker cluster for the left SN and right SN were 0.61 (0.05) and 0.59 (0.05) respectively.

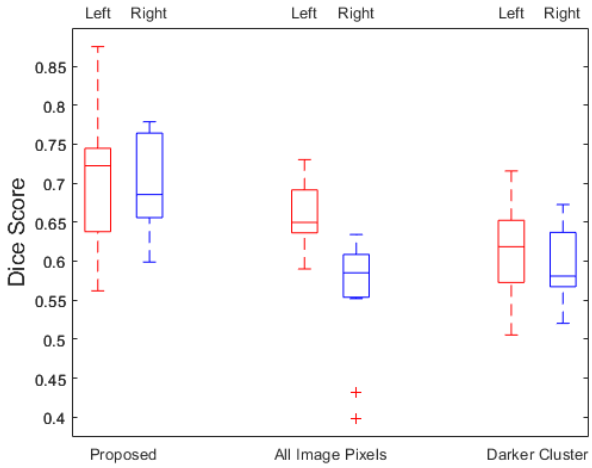


Figure 4.7: Boxplot of Dice scores of the proposed segmentation method, all image pixels, and the darker cluster for the left and right sides of SN.

We proposed an improved segmentation algorithm by estimating the threshold for accurate segmentation of SN using different samples of the image. To show the superiority of the proposed approach, we also estimated the threshold using all the pixels in the image which clearly could not segment SN effectively as shown in Figure 4.6. In fact, it covered many other irrelevant areas and could not contain the segmentation within the boundary of SN due to which the radiologists' gave a lower subjective score. Additionally, we also performed comparative analysis using the threshold of the darker cluster but it under-segmented SN as evident by the results in Figure 4.6.

The Dice score was used for the quantitative analysis in our work because it is the most commonly used metric to measure the fraction of spatial overlap between the binary regions for the evaluation of the segmentation result in the literature. The Dice score obtained for substantia nigra in this study was satisfactory.

In our study, the Dice score for the method based on all image pixels was penalized with a specific value following the expert clinicians' perspective since it fails to segment the midbrain nuclei within its boundary. The penalty was based on the subjective score provided by the radiologists' (Score 5: 0; Score 4: 0.07; Score 3: 0.14; Score 2: 0.21; Score 1: 0.28). The proposed method and the darker cluster could confine the segmentation and hence their Dice were not penalized.

Otsu thresholding (Otsu, 1979) was also used for the estimation of threshold in our proposed work because it is one of the most successful method for image segmentation based on minimizing within-class variance. Otsu thresholding was also performed on both cross-diagonal pixels and the remaining pixels. The threshold value estimated from Otsu thresholding was similar to *k*-means algorithm. Therefore, the segmentation result of Otsu method and our proposed method was similar/same. However, it should be noted that Otsu method is an exhaustive search algorithm that finds the global optimal threshold while *k*-means is a local optimal method. Moreover, *k*-means runs faster (D. J. Liu & Yu, 2009) than the Otsu method.

In the proposed methodology, the width of the cross-diagonals also affects the result. We dilated cross-diagonal pixels with an appropriate structuring element. In our work, we obtained the acceptable result with disk structuring element of radius *r*, where $2 \leq r \leq 4$. The structuring element of radius greater than this range results in an inefficient segmentation of SN. Additionally, we also assumed that the region of interest lies approximately in the center of the image. Hence, the target region was selected in such a way that the region of interest lie almost in the center of the image.

However, the proposed weighted thresholding method was challenged by the cases as shown in Figure 4.8. In these cases, the intensity of SN was very similar to those around it, especially in the upper part of SN. Because of this reason, the proposed method estimated the threshold value larger than it was desirable and as a result under-segmented SN. Additionally, the desirable SN segmentation obtained from the proposed method also leaves room for improvement, specifically, enclosed area of SN could be increased employing better techniques. Therefore, we aim to develop more accurate, robust, and efficient algorithm for the segmentation of midbrain nuclei and solve the challenges set by the proposed weighted thresholding method. Precisely, we focus on developing effective segmentation method utilizing level set method as it is a sound technique for evolving curves and surfaces, and is used in numerous applications as it is successful in capturing topological changes of objects in an automatic way.

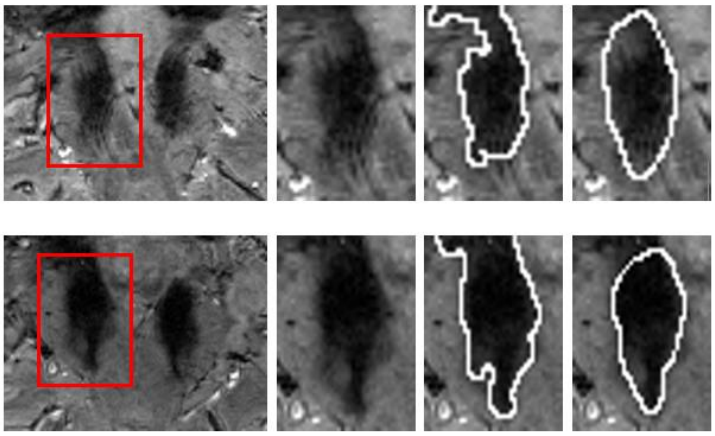


Figure 4.8: Results of the proposed segmentation method and ground truth segmentation. Column 1: Bounding box selected manually. The box is represented in red. Column 2: Area within the bounding box to initiate the segmentation. Column 3: Segmentation result of our proposed method. Column 4: Manual segmentation of SN. The segmentation is shown as a white contour.

4.6.2 Algorithm Performance (Level Set and Wavelet Transform)

The proposed segmentation method of combining level set and dual-tree complex wavelet transform for accurate segmentation of substantia nigra requires various parameters to be specified. We use the following parameters for the level set method for various MR images in our experiments: $\lambda_1 = 1.0$, $\lambda_2 = 2.0$, time step $\Delta t = 0.1$, $\mu = 1$, $v = 0.003 \times 255 \times 255$ and $\sigma = 10.0$. The level set parameters were defined after evaluation with several variations, especially the scale parameter σ.

The segmentation results of the proposed algorithm, comparison with different level set methods and manual segmentations of the substantia nigra are shown in Figure 4.9. The result shows that the proposed segmentation method produced good results and are closer to the manual segmentation. However, the RSF model and level set method based on the local intensity clustering property could not segment SN effectively and efficiently.

The proposed segmentation method received a higher subjective score from both radiologists' in comparison to the level set methods. The subjective evaluation scores for the sample scans presented in Figure 4.9 from both radiologists' are shown in Table 4.4 (RJK) and Table 4.5 (AL). The radiologists' liked the smooth result of the proposed methodology and stated that it enclosed better SN in comparison to the level set methods.

The inter-rater agreement between the subjective scores (ICC) is also reported in Table 4.6.

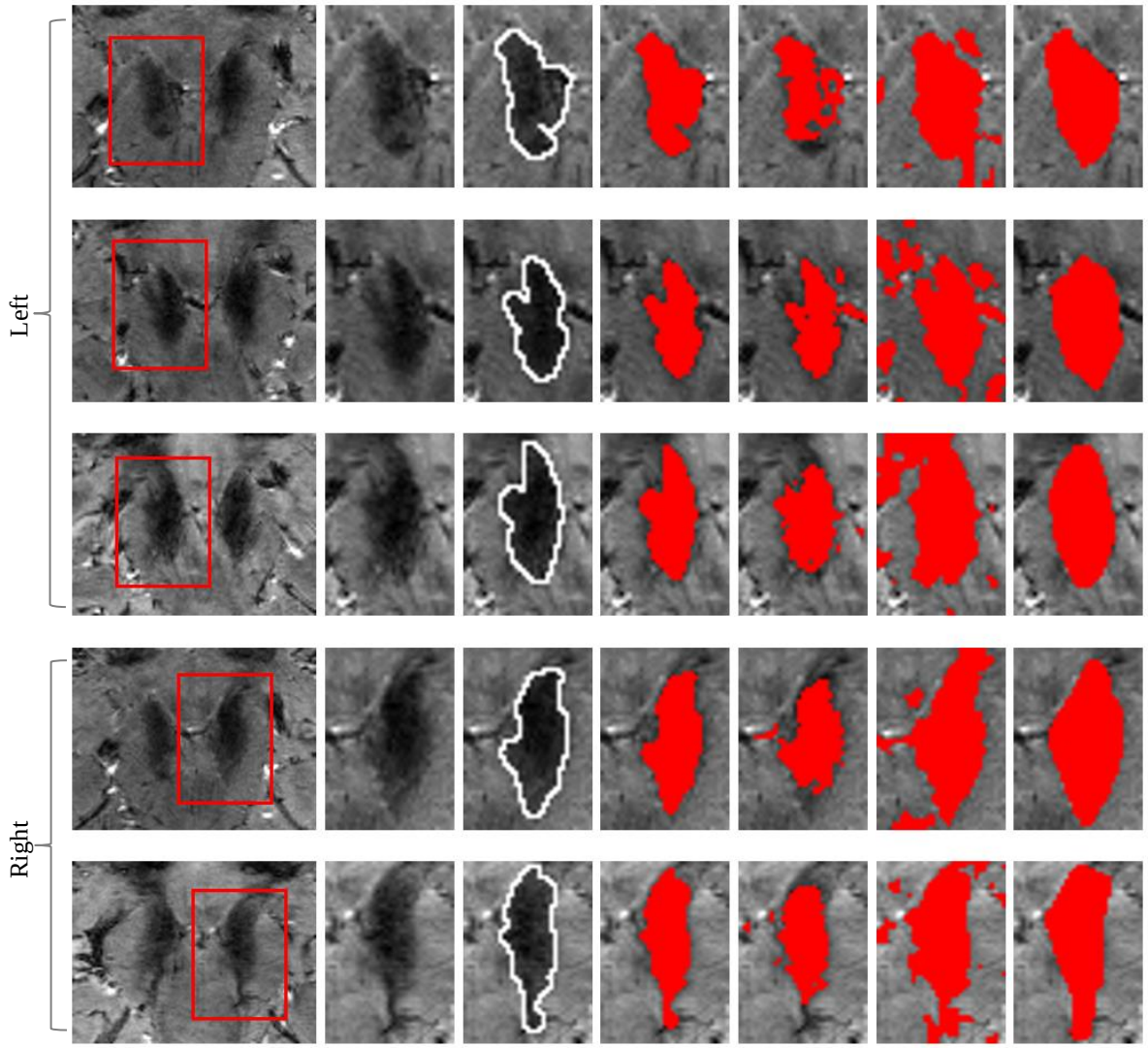


Figure 4.9: Results of the proposed segmentation method and comparison with different methods. Column 1: Bounding box selected manually. The box is represented in red. Column 2: Area within the bounding box to initiate the segmentation. Column 3: Segmentation result of our proposed method. The result of segmentation is shown as a white contour. Column 4: labeled mask of the proposed segmentation method is shown in red (SN). Column 5: Result of the level set method based on minimization of region-scalable fitting energy. Column 6: Result of the level set method in the presence of intensity inhomogeneities with application to MRI. Column 7: Manually labeled structures of SN.

Table 4.4: Subjective Evaluation Scores (Radiologist 1: RJK)

	Scan Number	Proposed Method	RSF	Local Intensity Clustering
Left	Scan 1	4	1	2
	Scan 2	3	2	1
	Scan 3	4	2	1
Right	Scan 4	4	3	1
	Scan 5	4	2	1

Score 1: very poor; Score 2: poor; Score 3: fair; Score 4: good; Score 5: very good

Table 4.5: Subjective Evaluation Scores (Radiologist 2: AL)

	Scan Number	Proposed Method	RSF	Local Intensity Clustering
Left	Scan 1	4	2	2
	Scan 2	3	2	1
	Scan 3	4	3	1
Right	Scan 4	4	3	2
	Scan 5	4	3	1

Score 1: very poor; Score 2: poor; Score 3: fair; Score 4: good; Score 5: very good

Table 4.6: Inter-rater agreement between the subjective scores

Method		
Proposed Method	0.82	Good
RSF	0.31	Poor
Local Intensity Clustering	0.4	Poor

The Dice score for the proposed method and RSF method was also calculated and is shown in Figure 4.10. The Dice score (mean (standard deviation)) of the left SN and right SN between the manual segmentation and the proposed method were 0.76 (0.06)

and 0.75 (0.06) respectively while the Dice for the RSF method for the left SN and right SN were 0.48 (0.06) and 0.49 (0.05) respectively.

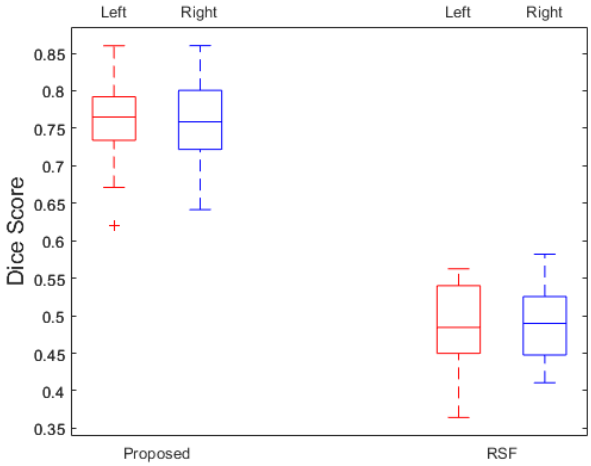


Figure 4.10: Boxplot of Dice scores of the proposed segmentation method and RSF method for the left and right sides of SN.

We proposed an effective method for the segmentation of smaller brain nuclei like substantia nigra by combining level set method and dual-tree complex wavelet transform. The level set method proposed in this work used the local image intensities rather than the global intensities which made it capable of handling images with intensity inhomogeneities common in medical images. Moreover, the level set method also avoided the need for time consuming and expensive re-initialization processes. However, level set method independently could not segment SN effectively. Hence, DT-CWT was used for accurate segmentation along with level set method. DT-CWT was capable of solving over-segmentation and under-segmentation, remove the undesired holes and smooth the rough output resulting from the level set method, and provided the desired segmentation of the midbrain nuclei.

The level set (C. Li et al., 2011; C. Li et al., 2008), an effective segmentation method was employed for comparative analysis in our work. However, the level set method was unable to segment SN effectively and efficiently, specifically due to the similarity in intensity profile with the adjacent areas, as well as unclear and weak boundaries of the smaller nuclei. Because of which, it constantly overestimated and underestimated the brain nuclei as shown in Figure 4.9. Moreover, the level set method resulted in over-segmentation of the midbrain nuclei in the form of holes, which did not make anatomical sense, according to expert clinicians' perspective.

The Dice score obtained for SN in this study is in agreement with the current literature (Haegelen et al., 2013; Kim et al., 2014; Visser et al., 2016; Yiming Xiao et al., 2012; Y. Xiao et al., 2014) although comparison between them does not make much sense because of the difference in acquisition parameters, manual segmentations, and the subjects involved in the experiments. Moreover, the duration of disease in PD patients also makes considerable difference in the study.

The Dice scores for the method based on the RSF was also penalized for over-segmentation including holes and its inability to confine the segmentation within the boundary of SN. The penalty was based on the subjective score provided by the radiologists' (Score 5: 0; Score 4: 0.07; Score 3: 0.14; Score 2: 0.21; Score 1: 0.28) like in earlier section. The Dice score for the level set method (C. Li et al., 2011) was not calculated because of the severe over-segmentation and under-segmentation present in the segmentation outputs.

As evident from the results, the Dice score obtained using the proposed methodology was better than the method based on the weighted thresholding which is shown in Figure 4.11. Similarly, the method based on the combination of level set and DT-CWT also addresses and solves the challenges of the weighted thresholding method as shown in the cases in Figure 4.12. The proposed method was capable of producing effective segmentation of SN in these cases as well. Hence, the proposed method based on level set and DT-CWT is more accurate, efficient, and robust than the weighted thresholding method.

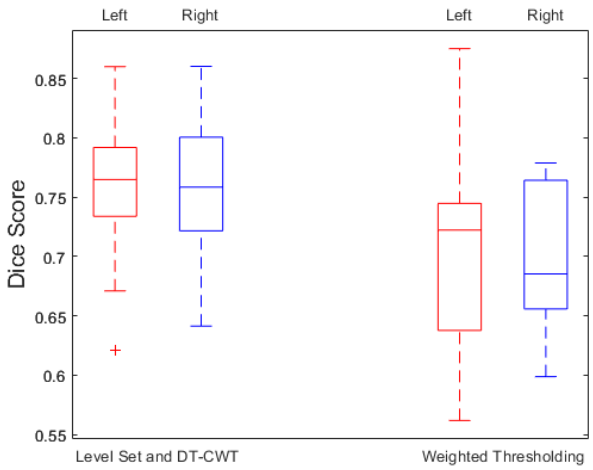


Figure 4.11: Boxplot of Dice scores of the Level Set and DT-CWT, and Weighted Thresholding for the left and right sides of SN.

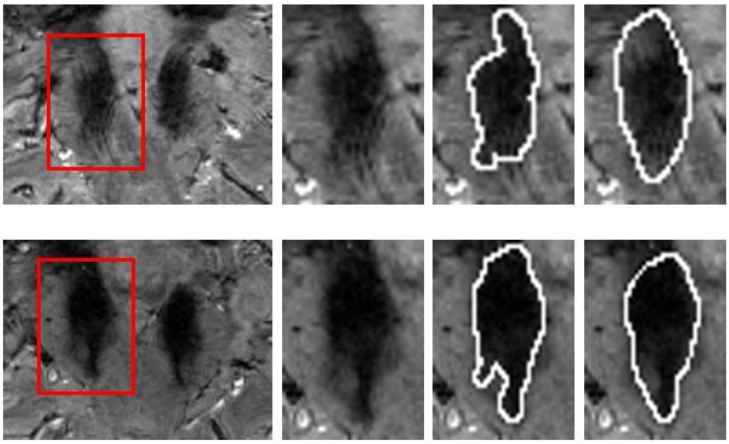


Figure 4.12: Results of the proposed segmentation method and ground truth segmentation. Column 1: Bounding box selected manually. The box is represented in red. Column 2: Area within the bounding box to initiate the segmentation. Column 3: Segmentation result of our proposed method. Column 4: Manual segmentation of SN. The segmentation is shown as a white contour.

4.7 Chapter Summary

This chapter presented improved algorithms for the segmentation of substantia nigra in T2*-weighted images. Substantia nigra was segmented using two approaches including weighted thresholding method and the method based on the combination of level set and dual-tree complex wavelet transform. The weighted thresholding approach used the weighted sum of threshold computed along the cross-diagonal pixels and the pixels outside the cross-diagonal pixels for estimating the final threshold value. On the other hand, the level set method used the local information of the input image while DT-CWT thresholded high frequency wavelet coefficients for the effective segmentation of SN.

The chapter also showed the effectiveness of the proposed methodologies through the experimental study using the images obtained from an online database consisting of healthy participants. Subjective evaluation performed by two radiologists' and quantitative analysis performed based on the manual segmentation performed by the radiologist further cemented the effectiveness of the algorithms. Substantia nigra segmentation performed using the proposed methods were closer to manual segmentation than the general thresholding methods and popular level set methods. However, method based on level set and DT-CWT overall proved to be superior to weighted thresholding method.

Despite reporting the highest mean Dice score for the proposed method based on level set and DT-CWT, one limitation is that the methodology was not completely automatic. In fact, the target region for the segmentation of substantia nigra was selected manually. Therefore, manual intervention was still required. Hence, the next chapter deals with selecting the target region using substantia nigra seeding to make the segmentation method completely automated along with developing more accurate, robust, and efficient algorithm for the segmentation of midbrain nuclei. The study also aims to focus on the segmentation of other smaller brain nuclei like red nucleus along with substantia nigra as it has been found to be involved in Parkinson's disease as well. We also extend our work to the new imaging technique, quantitative susceptibility mapping, which provides excellent contrast of the midbrain structures using our own database provided by the New Zealand Brain Research Institute consisting a large group of healthy and Parkinson's disease patients.

5 SEGMENTATION OF SUBSTANTIA NIGRA AND RED NUCLEUS USING QUANTITATIVE SUSCEPTIBILITY MAPPING IMAGES: PARKINSON'S DISEASE CHARACTERIZATION

Quantitative susceptibility mapping, an emerging magnetic resonance imaging technique, allows high-resolution quantification of magnetic susceptibility, producing images of the small midbrain structures with excellent contrast (de Rochefort et al., 2010; Schweser, Sommer, Deistung, & Reichenbach, 2012). QSM provides an improved contrast compared to T2*-weighted images (Alkemade et al., 2017; T. Liu, Eskreis-Winkler, et al., 2013). The blooming artifacts present in T2*-weighted images also makes tissue border definition difficult, while QSM eliminates the problem of these blooming artifacts (Y. Wang et al., 2017). Moreover, QSM measures the local property

which is crucially different from the nonlocal property of the GRE magnitude T2*-weighted imaging.

Iron is highly paramagnetic and therefore induces a large local magnetic distortion, which is detectable with QSM *in vivo* (Bilgic, Pfefferbaum, Rohlfing, Sullivan, & Adalsteinsson, 2012; Deistung et al., 2013). Recent work employing QSM has reported increased susceptibility in PD relative to controls in basal ganglia structures, including the substantia nigra and red nucleus, and even cortical regions (Chen et al., 2019; Thomas et al., 2020; Uchida et al., 2019).

In order to effectively investigate the relationship between iron accumulation and PD, we developed and evaluated an improved segmentation algorithm that uses information from QSM images to segment both SN and RN. The segmentation algorithm proposed in this chapter is an enhancement to the method described in Chapter 4 with level set and wavelet transform. The methodology has been improved with additional substantia nigra seeding, contrast enhancement and watershed transform steps to effectively segment and separate SN and RN. The experiments in this chapter were performed using the database provided by New Zealand Brain Research Institute utilizing QSM images from 20 healthy controls and 20 PD participants. Of the 20 PD participants, seven had normal cognition, eight had mild cognitive impairment, and five had dementia. After comparing algorithm-derived segmentations to 'gold standard' segments manually identified by experienced radiologists', we investigated differences in QSM values between PD and controls, and associations with cognitive and motor dysfunction in PD, from both gold standard and algorithm-derived SN and RN segmentations.

5.1 Segmentation Method

The segmentation method proposed in this chapter consists of different steps including substantia nigra seeding, contrast enhancement, level set method, wavelet transform, and watershed transform. The level set method and wavelet transform proposed in this chapter is same as the one described in the previous chapter. The flowchart of the proposed segmentation method is shown in Figure 5.1.

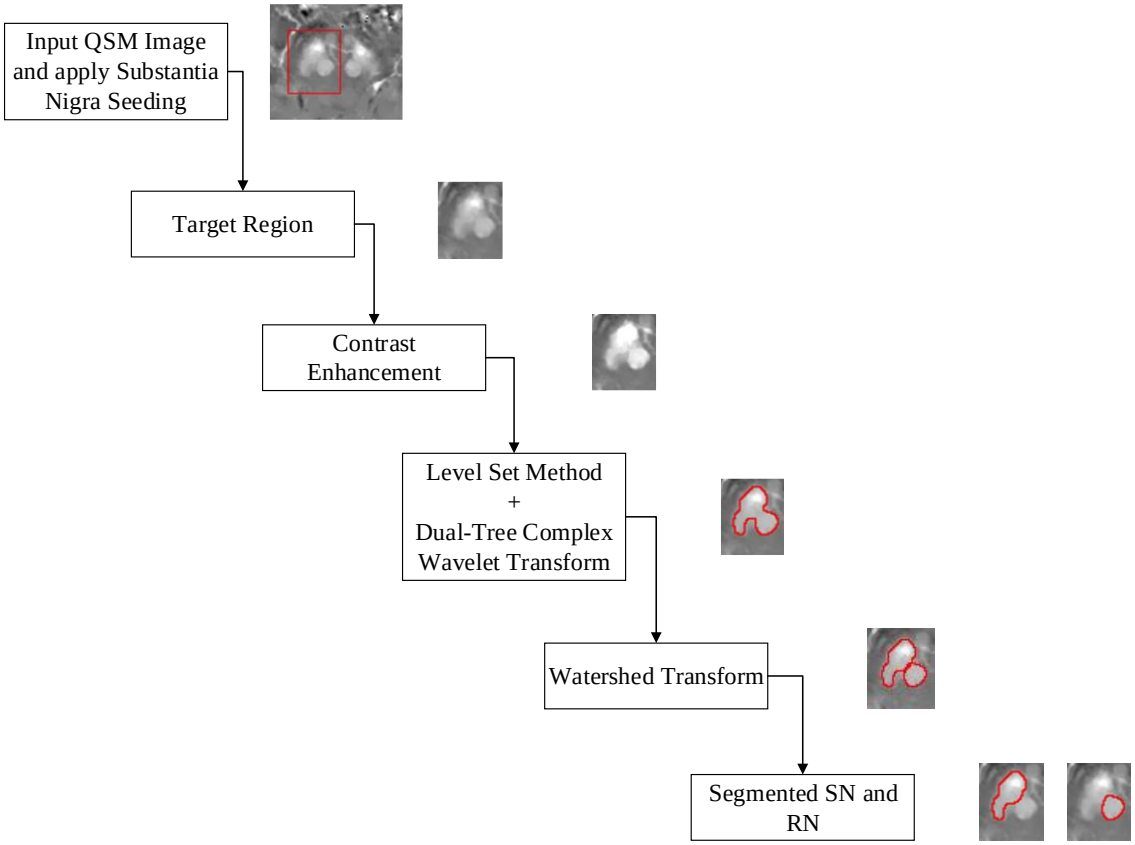


Figure 5.1: Flowchart of the proposed method

5.1.1 Substantia Nigra Seeding

On the normalized, study-specific average QSM image, a single voxel was manually placed within the SN on each of six slices, creating left and right SN seed images in Montreal Neurological Institute space. These MNI-space seeds were then inverse-warped into QSM subject space using the inverse deformation fields produced during the structural T1-weighted image processing. This produced seeds within the left and right SN in each individual; these seeds were used to initiate the SN and RN segmentations by creating a bounding box within the target region.

5.1.2 Contrast Enhancement

Contrast enhancement proposed in this thesis was based on local and global gray level information of the input image. Local gray level information was obtained by computing the mean of each non-overlapping 3×3 sub-image using mean filter while global gray level information was estimated by calculating the mean of the entire image. The difference between the global gray level mean and the local gray level mean

calculated over a neighbourhood of each pixels helps in identifying the gray level changes which depends on image characteristics in that neighbourhood (Khan, Arya, & Pattanaik, 2013; Singh, Yadav, & Singh, 2016).

Let $A(x,y)$ be an input image. The global mean of this image can be calculated from entire image sample values as follows:

$$g_{mean} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} A(x,y) \quad \textbf{5.1}$$

Where *M, N* denote the image size (width, height respectively) in pixels.

The local mean was calculated on each non-overlapping 3×3 sub-image using an average filter. Let the sub-image of specified size be centered on (x,y). The contrast enhancement was performed on each sub-image as follows:

$$I(x+i,y+j) = A(x+i,y+j) + [l_{mean}(x,y) - g_{mean}] \quad \textbf{5.2}$$

Where $i,j = -1,0,1$, $A(x+i,y+j)$ is the input image, $l_{mean}(x,y)$ is the local mean at (x,y), and $I(x+i,y+j)$ is the contrast enhanced sub-image. The overall contrast-enhanced image $I(x,y)$ was computed by repeating the process in equation (5.2) for all the sub-images.

5.1.3 Level Set Method

The level set method used in this chapter was described in Section 4.2.1 in Chapter 4.

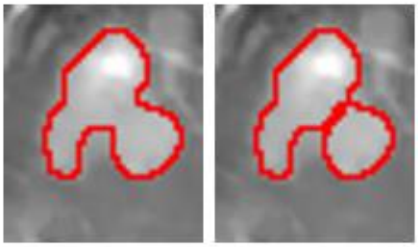
5.1.4 Wavelet Transform

In the research work in this thesis, we used a dual-tree complex wavelet transform for smoothing which was discussed in detail in Section 4.2.2.1 in previous chapter.

5.1.5 Watershed Transformation

The watershed transform (Roerdink & Meijster, 2000; Vincent & Soille, 1991) is a popular segmentation method that has been widely used in medical image segmentation because of its numerous advantages. It is a simple, fast, intuitive method, and produces a complete division of the image into distinct regions even when the contrast is poor,

thus avoiding the need for any kind of post-processing work such as contour joining as shown in Figure 5.2.



(a) (b)

Figure 5.2: (a) SN and RN contour merged together (b) Separation of SN and RN using watershed transform.

Distance transform (Borgefors, 1986) was computed before applying the watershed transform. The distance transform labels each pixel of the image whose value correspond to the distance to the nearest feature pixel. The watershed transform subsequently performed on the distance transformed image results in over-segmentation—the main drawback of this algorithm—because of the presence of several local minima. Therefore, an extended-minima transform was used to modify the distance-transformed image for accurate and effective segmentation.

5.1.5.1 Extended-Minima Transform

The extended-minima transform $EMIN$ are defined as the regional minima of the corresponding H -minima transformation (Cheng & Rajapakse, 2009; Soille, 2003).

$$EMIN_h(f) = RMIN[HMIN_h(f)] \quad \textbf{5.3}$$

Where,

$$HMIN_h(f) = R_f^{\varepsilon}(f+h) \quad \textbf{5.4}$$

Is the H -minima transformation and is achieved by performing the reconstruction by erosion of f with respect to $f+h$. H -minima transform suppresses all minima whose depth is below or equal to the given h.

5.1.6 Implementation

The bounding box for substantia nigra and red nucleus segmentation was obtained by substantia nigra seeding. The contrast within the bounding box can be improved for better and effective segmentation. Therefore, contrast enhancement was performed on the values within the bounding box using equation (5.2) for all the sub-images. The contrast enhanced image helped in directing the evolution of level set function towards the desired direction.

The boundaries of the segmentations were obtained with the level set method while DT-CWT smoothened the jagged output and resolved over-segmentation and under-segmentation resulting from level set method. The implementation detail of level set method and DT-CWT can be found in the previous chapter. But, it was observed that the combination of the level set method and DT-CWT was still not capable of separating SN and RN accurately and efficiently. Therefore, distance transformation was performed on the binary image resulting from the previous steps. Extended-minima transform was used to suppress undesired local minima and then the distance transformed image was modified so that no minima appeared at the filtered-out location. Watershed transform performed on this modified distance transformed image helped separating and segmenting SN and RN effectively and efficiently.

5.2 Qualitative Assessment of Segmentations

The substantia nigra and red nucleus in this chapter were segmented using three different methods: (1) our proposed method, (2) a level set method based on minimization of region scalable fitting energy for segmentation of images with intensity inhomogeneities proposed by Li et al. (C. Li et al., 2008), and (3) a different level set method proposed by Li et al. (C. Li et al., 2011) that is based on the local intensity clustering and is capable of handling images with intensity inhomogeneities.

Two experienced radiologists' (AL and MH) separately performed subjective evaluation of the segmentations produced by the three methods described above. Subjective evaluation was performed blinded to method, group, and demographic information. A score ranging from 1-5 was used for the subjective assessment of the segmentations.

5.3 Quantitative Assessment: Expert ('Gold standard') Manual Segmentations

An experienced radiologist (AL) performed the manual segmentations of SN and RN on all 40 subjects. A second experienced radiologist (MH) verified the manual segmentations drawn by the first radiologist; in the case of disagreement, the manual segmentation was adjusted until both radiologists' were in agreement. Radiologists' were blinded to all clinical and demographic characteristics. These manual segmentations were considered as the gold standard for further analyses.

Dice score (given in equation (3.1)), a global measure of segmentation quality was used for quantitative assessment in this chapter to estimate the degree of overlap between the two segmentations.

5.4 Experimental Results

We used the following parameters for the level set method in our experiments: $\lambda_1 = 1.0$, $\lambda_2 = 2.0$, time step $\Delta t = 0.1$, $\mu = 1$, $v = 0.003 \times 255 \times 255$ and $\sigma = 10.0$. The level set parameters used in this chapter were obtained during the experimental studies using T2*-weighted images in Chapter 4.

5.4.1 Algorithm Performance

The methodology proposed in this chapter produced good segmentation results (Figure 5.3). The proposed segmentation method received a higher subjective score from both radiologists'. The alternative level set methods did not segment or separate SN and RN effectively. The subjective score for the sample scans presented in Figure 5.3 is shown in Table 5.1 (AL) and Table 5.2 (MH). The inter-rater agreement between the subjective scores (ICC) is also displayed in Table 5.3.

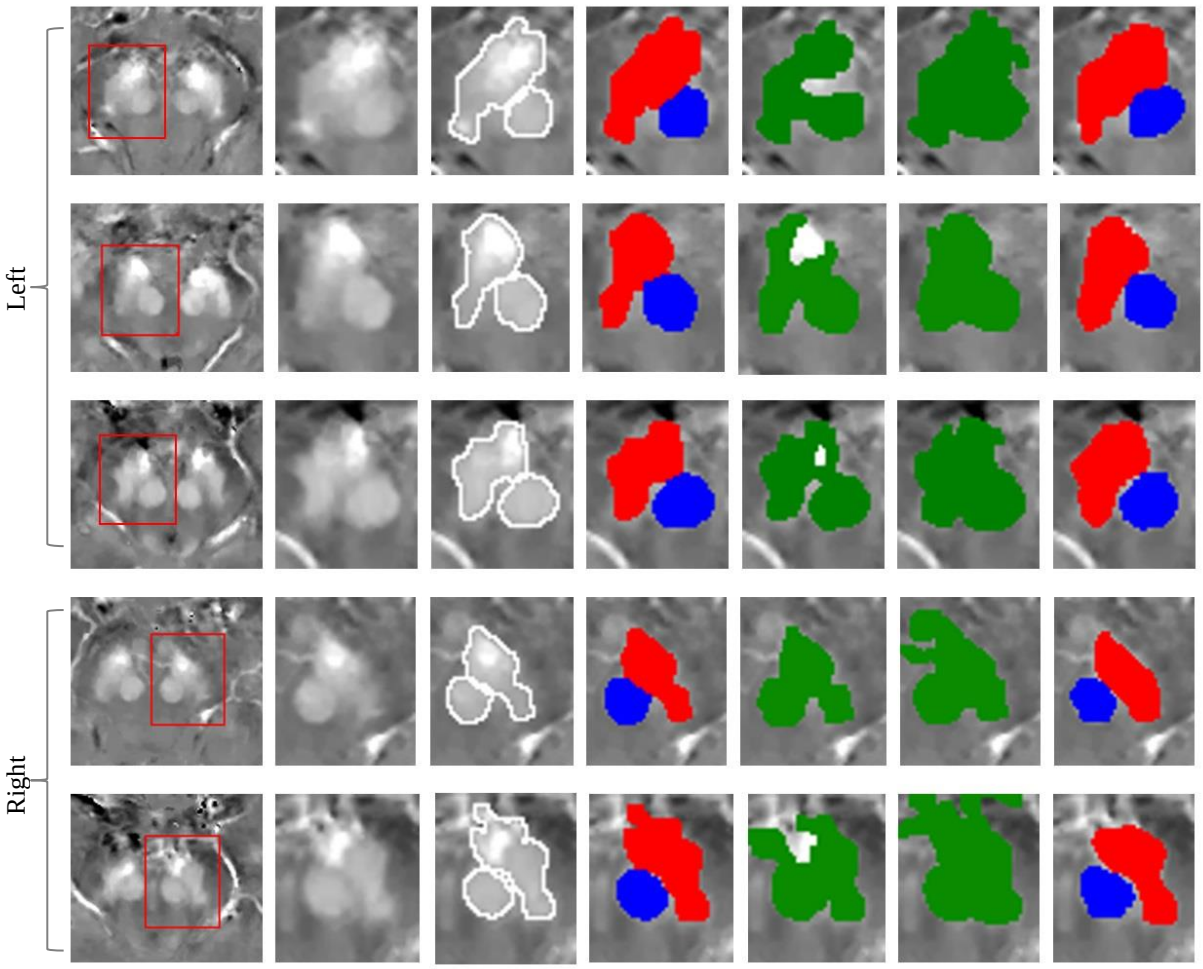


Figure 5.3: Results of the proposed segmentation method and comparison with different methods. Column 1: Bounding box obtained by substantia nigra seeding. The box is represented in red. Column 2: Area within the bounding box to initiate the segmentation. Column 3: Segmentation result of our proposed method. The result of automated segmentation is shown as a white contour. Column 4: labeled mask of the proposed segmentation method is shown in red (SN) and blue (RN). Column 5: Result of the level set method based on minimization of region-scalable fitting energy. Column 6: Result of the level set method in the presence of intensity inhomogeneities with application to MRI. SN and RN both are labeled as green since level set method cannot separate them. Column 7: Manually labeled structures of SN (red) and RN (blue).

Table 5.1: Subjective Evaluation Scores (Radiologist 1: AL)

	Scan Number	Proposed Method	RSF	Local Intensity Clustering
Left	Scan 1	5	2	1
	Scan 2	5	1	1
	Scan 3	5	2	1
Right	Scan 4	4	2	1
	Scan 5	4	1	1

Score 1: very poor; Score 2: poor; Score 3: fair; Score 4: good; Score 5: very good

Table 5.2: Subjective Evaluation Scores (Radiologist 2: MH)

	Scan Number	Proposed Method	RSF	Local Intensity Clustering
Left	Scan 1	5	2	1
	Scan 2	4	1	1
	Scan 3	5	1	1
Right	Scan 4	4	2	1
	Scan 5	4	1	1

Score 1: very poor; Score 2: poor; Score 3: fair; Score 4: good; Score 5: very good

Table 5.3: Inter-rater agreement between the subjective scores

Method		
Proposed Method	0.77	Good
RSF	0.80	Good
Local Intensity Clustering	0.87	Good

In the quantitative analysis, our proposed method produced good Dice scores (Figure 5.4). The Dice score (mean (standard deviation)) of the left SN and right SN between the gold standard segmentation and the proposed method were 0.75 (0.09) and 0.76 (0.07), respectively; left RN and right RN were 0.81 (0.07) and 0.77 (0.07), respectively.

Initially, we performed the experimental analysis on 20 subjects (including both healthy controls and PD participants) using the expert manual segmentation from single radiologist. The Dice score of the left SN and right SN between the gold standard segmentation and the proposed method were 0.77 (0.09) and 0.78 (0.07), respectively; left RN and right RN were 0.80 (0.08) and 0.77 (0.08), respectively.

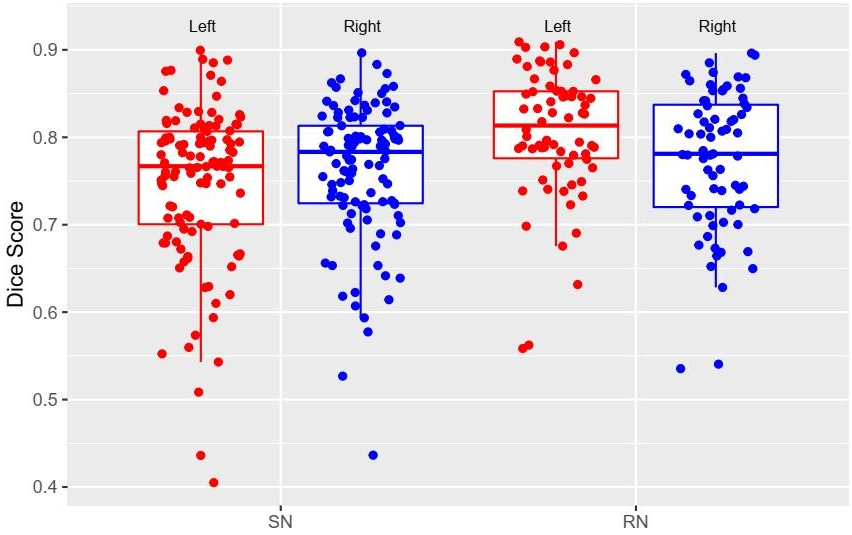


Figure 5.4: Boxplot of Dice scores of the proposed segmentation method for the left and right sides of SN and RN. Dots are Dice scores for all the slices of images used in the experiment from all 40 subjects.

5.5 Relationship between Gold Standard and Proposed Method QSM Values

Median QSM values were extracted from gold standard SN and RN segments, as well as from segments produced with our proposed method (right and left were extracted independently). Pearson correlation coefficients were used to assess the correlation between QSM values extracted from the two methods; reliability was measured using the intraclass correlation coefficient type (A,1) and (C,1) (McGraw & Wong, 1996).

5.6 Parkinson's Disease-Related Differences in Susceptibility

Median QSM values from both 'gold standard' and our proposed method SN and RN ROIs were compared across group, global cognitive ability, and motor impairment using Bayesian regression models. Bayesian models were fitted using the "brms" (v2.9.0)

package (Bürkner, 2017) in R (v3.6.3). In each model, four chains with 5,000 iterations each were used to generate the posterior sample. Given previous research, we examined hypotheses of increased QSM in PD relative to controls, a negative association with global cognitive ability, and a positive association with motor impairment. Parameter estimates, 95% uncertainty intervals, and probabilities of hypotheses being true are presented.

(1) Group differences: We first tested for evidence of varying susceptibility (QSM value) between PD and healthy control groups. For each segmentation method (gold standard/proposed method), we modelled QSM as a function of age, sex, group (PD/control), side (left or right), ROI (SN/RN), and the interactions age-by-ROI (allowing the relationship between age and QSM to vary by ROI), group-by-ROI (allowing the difference in QSM between PD and controls to vary between ROI), and group-by-side (allowing the PD effect to vary between left and right side), with an intercept varying by subject, side, and ROI.

(2) Cognitive ability: In PD participants only, we modelled QSM as a function of age, sex, side (left or right), ROI (SN/RN), global cognitive ability (aggregate cognitive z score), and the interactions age-by-ROI, cognitive score-by-ROI, and cognitive score-by-side, with an intercept varying by subject, side, and ROI.

(3) Motor impairment: In PD participants only, we modelled QSM as a function of age, sex, side (left or right), ROI (SN/RN), UPDRS III, and the interactions age-by-ROI, UPDRS III-by-ROI, and UPDRS III-by-side, with an intercept varying by subject, side, and ROI.

5.7 Application to Parkinson's disease

The demographic and clinical data used in this chapter are summarized in Table 5.4. Figure 5.5 presents the relationship between gold standard and proposed method QSM values in both SN and RN. Slope (m) and Pearson correlation coefficients (r), with 95% confidence intervals, were as follows: left SN: m = 0.94 [0.82, 1.06], r = 0.93 [0.87, 0.96], right SN: m = 0.92 [0.76, 1.07], r = 0.89 [0.80, 0.94], left RN: m = 0.92 [0.83, 1.00], r = 0.96 [0.93, 0.98], right RN: m = 0.98 [0.87, 1.09], r = 0.95 [0.90, 0.97]. ICC showed moderate-good agreement and excellent consistency (Table 5.5).

Table 5.4: Demographic and Clinical Information

Measure	Control (n=20)	Parkinson's Disease (n=20)
Sex (M:F)	14:6	14:6
Age	72 (8)	74 (6)
Global cognitive ability	0.8 (0.4)	-0.8 (1.0)
MoCA	27 (2)	22 (4)
UPDRS-III	-	41 (12)
Duration of diagnosis (years)	-	8 (5)

Data are mean (Standard Deviation). Sex is reported as a ratio. M : Male; F : Female. MoCA: Montreal Cognitive Assessment; UPDRS-III: Unified Parkinson's Disease Rating Scale Part 3. Global cognitive ability was calculated as an aggregate of standardized scores across five cognitive domains.

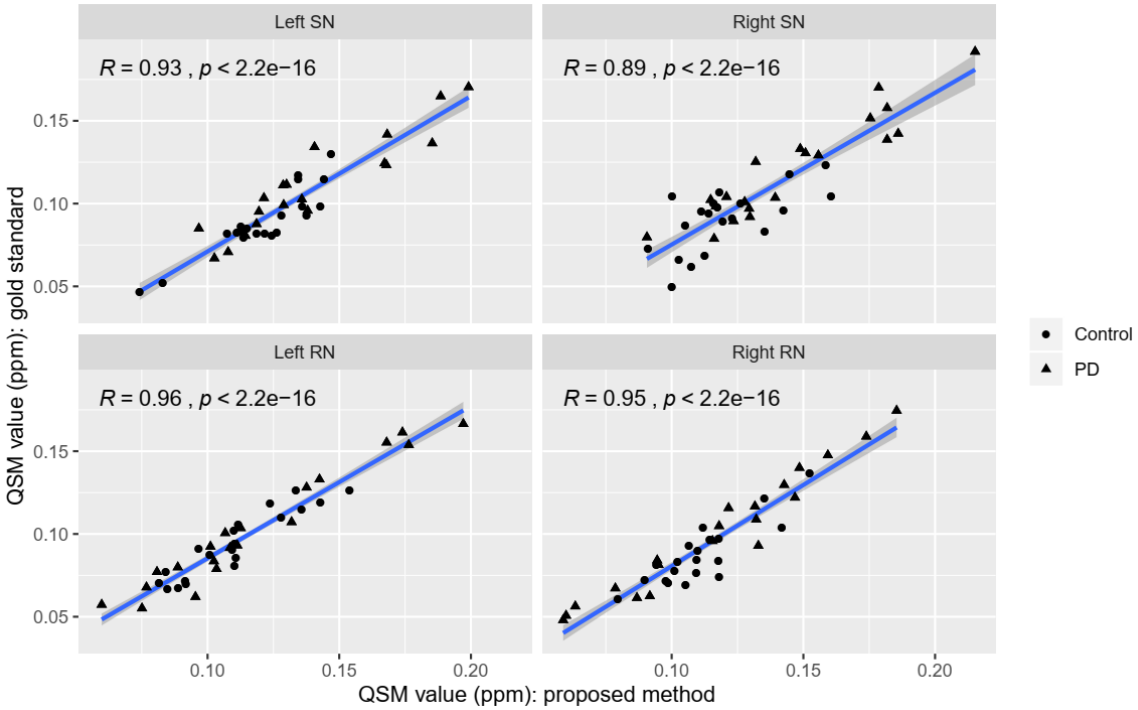


Figure 5.5: Relationship between gold standard QSM values and proposed method QSM values. R = Pearson Correlation Coefficient, p = p value. The line of best fit is displayed in blue, with 95% confidence intervals shaded in gray.

Table 5.5: Intraclass correlation coefficient for agreement and consistency

	ICC(A,1): Left	ICC(A,1): Right	ICC(C,1): Left	ICC(C,1): Right
SN	0.56 [-0.04, 0.86]	0.62 [-0.08, 0.88]	0.93 [0.87, 0.96]	0.89 [0.80, 0.94]
RN	0.84 [-0.04, 0.96]	0.77 [-0.06, 0.94]	0.96 [0.93, 0.98]	0.94 [0.90, 0.97]

Values are ICC [95% CI]. ICC(A,1) = Intraclass correlation coefficient for agreement, ICC(C,1) = Intraclass correlation coefficient for consistency. CI: Confidence interval.

5.7.1 Group Differences

Table 5.6 and Figure 5.6 report the results from comparisons between healthy control and PD groups in SN and RN for both the gold standard segmentations and those from the proposed method. We observed moderate evidence of increased QSM values in the SN in the PD group relative to controls (probability P = 98% (proposed) and P = 99% (manual segmentation) of the PD group having higher QSM values than the control group), with weak evidence within the RN (P = 75% (proposed) and P = 84% (manual). Furthermore, we found moderate evidence for an association with age in the SN (probability P = 92% of a positive association, proposed method), but no evidence in the RN (P = 29%), and no evidence of a difference between anatomic side (probability P = 48% probability of right side having greater QSM values).

Table 5.6: QSM differences between Parkinson's disease and control participants

	Gold standard		Proposed method	
Region	SN	RN	SN	RN
Control group (ppb)	85 [68, 101]	84 [67, 101]	115 [98, 132]	104 [86, 121]
PD vs controls (ppb)	20 [6.1, 34]	8.7 [-5.7, 23]	18 [3.7, 33]	5.9 [-8.9, 20]
Posterior probability of PD > controls	99%	84%	98%	75%

ppb = parts per billion. Posterior probability refers to the probability of PD being greater than controls.

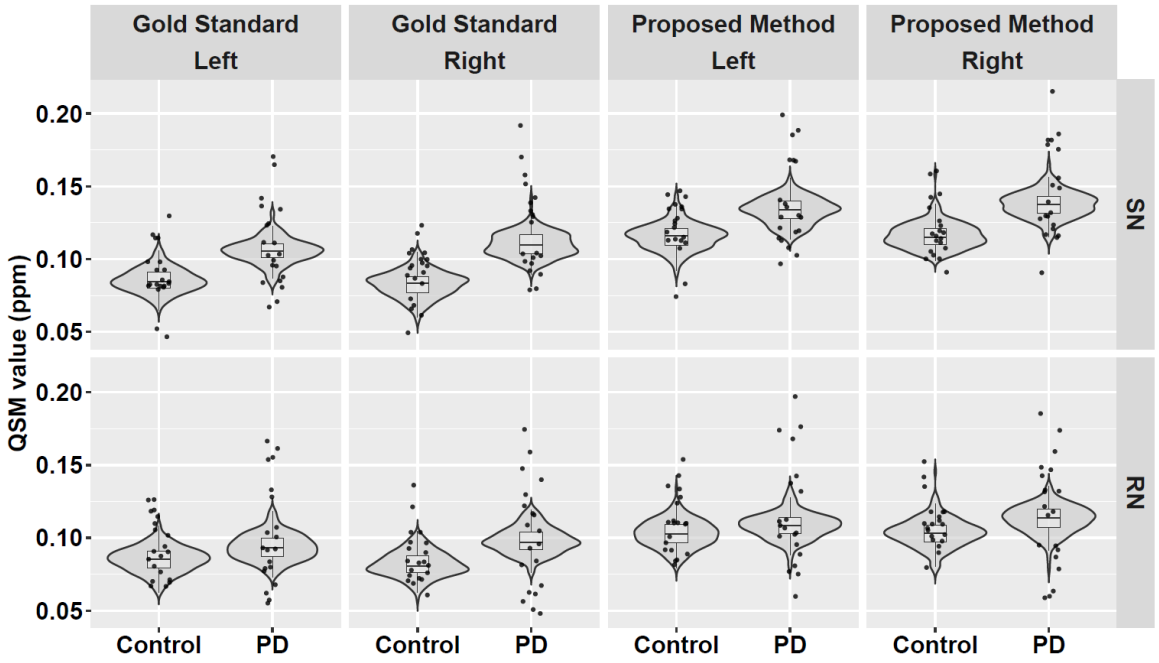


Figure 5.6: QSM values of left and right sides of SN and RN between healthy control and PD. Dots are raw QSM values. Violin plots in gray and boxplots represent the distribution of QSM values adjusted for covariates from 200 draws from the posterior probability distribution.

5.7.2 Cognitive Impairment, Motor Impairment, and QSM:

Table 5.7, Figure 5.7, and Figure 5.8 report the relationship between QSM, global cognitive ability, and motor impairment. We observed moderate evidence of a negative relationship between QSM and global cognitive z score in SN (94% (proposed) and 87% (manual segmentation) probability of a negative association between QSM and global cognitive ability). There was no evidence of association within RN (66% (proposed) and 61% (manual segmentation)) and no evidence of a difference in association between left and right (-0.11 ppb/cognitive z point [-15, 15]). There was moderate evidence for a positive association with motor impairment in SN (87% (proposed) and 95% (manual segmentation)) and within RN (96% (proposed) and 98% (manual segmentation)).

Table 5.7: Associations between QSM, global cognitive ability, and motor impairment

	Gold standard		Proposed method	
Region	SN	RN	SN	RN
Global cognitive ability (ppb/point)	-10 [-25, 5]	-2 [-17, 12]	-14 [-29, 0.8]	-4 [-18 , 12]
Posterior probability of a negative association	87%	61%	94%	66%
UPDRS-III (ppb/point)	13 [-0.2, 27]	18 [4, 32]	10 [-5, 25]	17 [2, 31]
Posterior probability of a positive association	95%	98%	87%	96%

The reported value [95% uncertainty interval] is the slope of the association: parts per billion per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

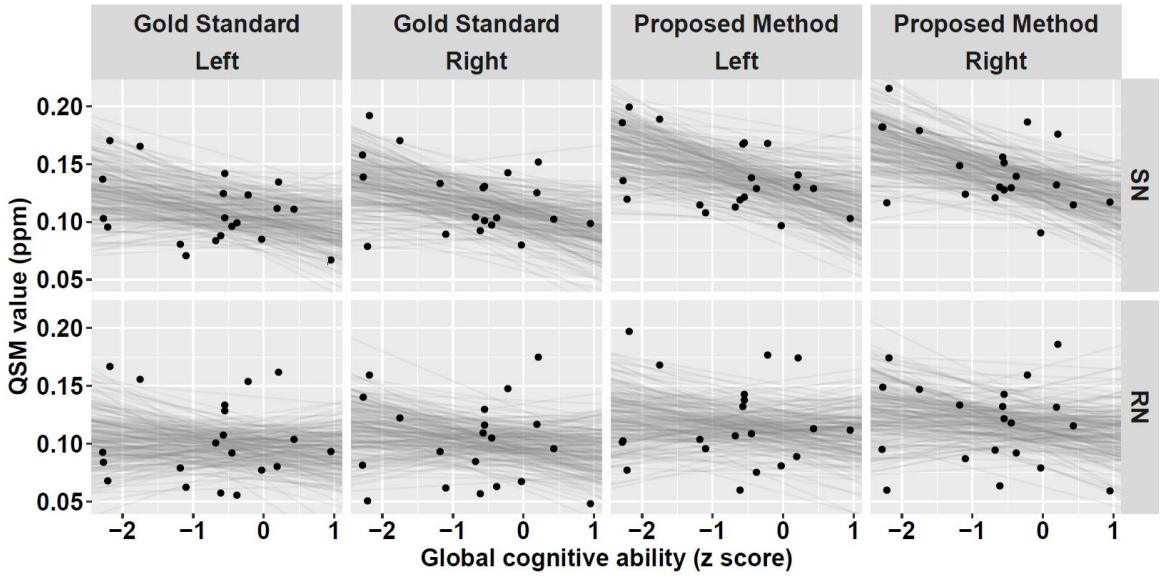


Figure 5.7: Correlations between QSM values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

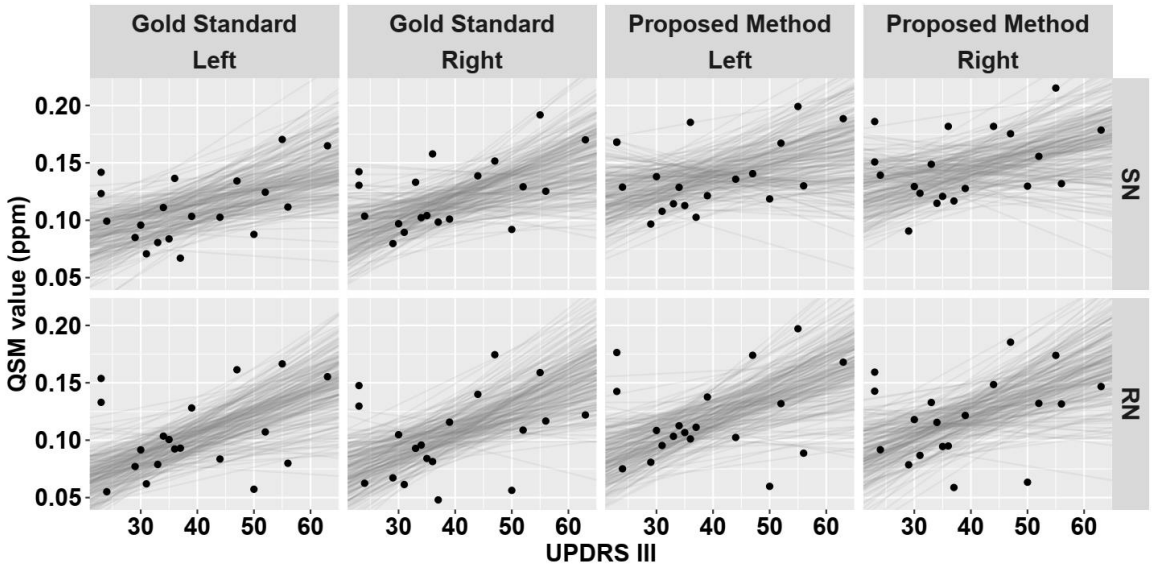


Figure 5.8: Correlations between QSM values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

5.8 Discussion

This chapter of the thesis proposed an effective automated method for SN and RN segmentation by combining substantia nigra seeding, contrast enhancement, level set method, dual-tree complex wavelet transform and a watershed transform. Segmentations from our proposed method were compared to radiologist-drawn 'gold standard' segmentations. The extracted QSM values were tested in a cohort of 20 healthy and 20 PD participants; consistent results between the two methods suggests our proposed segmentation algorithm may be used in the future with confidence.

The proposed automated segmentation method of combining various techniques progressed in the following manner. Substantia nigra seeding helped to obtain the bounding box within the target region containing SN and RN to initiate the segmentation. The contrast enhancement technique, which is based on the local and global mean, greatly improved the contrast of the region within the bounding box, which in turn helped in enhancing the contrast of the ROI. Specifically, contrast enhancement helps to increase the accuracy and effectiveness of segmentation as well as directing the evolution of level set function towards the desired direction. In spite of all the advantages of the proposed level set method mentioned in the earlier chapter, it still resulted in jagged output, over-segmentation, under-segmentation, and did not separate SN from RN

in some slices. Therefore, DT-CWT was used while evolving the level set function to address these issues, producing smooth output, and resolving over-segmentation and under-segmentation. However, the midbrain nuclei were still not effectively separated with this approach. Hence, a watershed transform based on the extended-minima transform was introduced to effectively separate and segment SN and RN. The combination of the above techniques resulted in the best segmentation that adequately matched expert manual segmentation.

The level set (C. Li et al., 2011; C. Li et al., 2008), a classical and effective segmentation method based on the change of gray level, was employed for comparative analysis in this study because it is frequently used in medical imaging. However, the level set method was unable to separate and segment SN and RN effectively and efficiently for the reasons similar to the ones mentioned in the previous chapter.

In our work, quantitative assessment was performed using Dice score. The Dice score obtained in this study is comparable to acceptable levels in the current literature on SN and RN (Guo et al., 2018; Visser et al., 2016). Such comparisons must be made with care, however, as image acquisition parameters, quality of manual segmentations, and specific sample heterogeneity may all influence this metric. For example, the partial volume effect and the absolute faint contrast of the midbrain nuclei present in some slices of few subjects limit segmentation accuracy and may reduce the overall Dice score. The Dice score for the level set methods was not calculated as these comparison methods were not able to separate the smaller midbrain structures. While we have taken the expert manual segmentation as the 'gold standard', there is a certain amount of uncertainty associated with these segments, as with any manual segmentation. However, we attempted to minimize this uncertainty by requiring agreement between two experts' which resulted in good reliability.

Additionally, we also performed pre-processing using contrast enhancement presented in this chapter to analyse the overall performance of segmentation method based on weighted thresholding introduced in Chapter 4. The results were found to be improved with additional pre-processing step. However, the segmentation based on level set method still had better performance in comparison.

In the current work, we applied our improved segmentation method to a cohort of PD patients. Quantitative susceptibility mapping allows in vivo investigation of iron accumulation in brain tissue, and recent work suggests increased iron deposition in SN in

PD (An et al., 2018; Barbosa et al., 2015; Du et al., 2016; Murakami et al., 2015). Our findings, of moderate evidence of increased QSM values in PD in the SN and in the RN, are consistent with the expected increase in iron accumulation in PD patients. QSM values with SN and RN exhibited moderate evidence of an association with motor severity in PD. This further suggests that iron accumulation may continue as the disease progresses, providing an imaging marker that tracks motor impairment. While pathology within the SN, and RN to a lesser extent, is generally associated with early disease and the emergence of motor impairments in PD, we also showed a moderate association between QSM values within the SN and global cognitive ability. A widespread pattern of abnormal iron deposition in both cortical and subcortical regions may also underlie cognitive impairments (Thomas et al., 2020; Uchida et al., 2019). Interestingly, these two recent studies (identifying SN manually or atlas-based) did not identify correlations with a global cognitive score (the Montreal Cognitive Assessment) or the UPDRS-III. The combination of improved segmentation of the SN using the proposed method along with more comprehensive neuropsychological assessments may help explain our moderate associations.

Our study has some limitations. First, the proposed algorithm is primarily based on the intensity and texture of the image. Therefore, the performance of the algorithm drops when the contrast is poor. Thus, better imaging modalities providing adequate contrast of the midbrain nuclei could be used for improved performance and segmentation accuracy. Second, the sample size used in the study for both groups was relatively small, which may not capture the full range of QSM and SN variability during PD progression. Third, we only considered the entire SN, aggregating pars compacta, pars reticulata, and dorsal and ventral sub-regions into the full SN. While dopaminergic cell death is prominent within the pars compacta in PD, studies investigating iron accumulation have reported increase in both the pars compacta and pars reticulata (Du et al., 2018). A recent study (Bergsland et al., 2019), assessed iron separately in the ventral and dorsal levels of the SN in PD and specifically found significant iron increases in the ventral SN of the PD patients relative to healthy controls. While the specificity of iron accumulation in specific sub-regions of the SN is of interest and should be explored in the future, we focused on segmentation of the whole SN given its small size, relatively large imaging voxels, and lack of obvious contrast within the QSM images. Thus, we are unable to comment on any differences across the differing SN

sub-regions, however future extension of the current segmentation algorithm should allow identification of SN sub-regions.

5.9 Chapter Summary

This chapter presented an improved automated segmentation method for the midbrain structures including substantia nigra and red nucleus from QSM images by combining substantia nigra seeding, contrast enhancement, level set method, dual-tree complex wavelet transform, and watershed transform. The experimental analysis were conducted using the QSM images obtained from the New Zealand Brain Research Institute consisting of 20 healthy controls and 20 PD participants. The proposed method produced high quality segmentations of SN and RN, validated against expert manual segmentations, and facilitates *in vivo* assessment of midbrain nuclei iron content. The study revealed the significant correlation between QSM values extracted from both gold standard and the proposed segmentation method for the left and right SN and RN.

This chapter also presented characterization of Parkinson's disease using the segmentations of SN and RN. The study showed increased QSM values in SN in PD relative to controls, moderate evidence of association between SN QSM values with global cognitive ability, and moderate evidence of association between SN and RN QSM values with motor impairment; these results were comparable between both our proposed segmentation and expert manual segmentation. The work presented in this chapter suggested that an improved midbrain segmentation algorithm may be useful for monitoring iron-related disease severity in Parkinson's disease. Having validated the proposed SN and RN segmentation method, we now have the opportunity to apply this technique to a large longitudinal cohort of PD patients and controls, to allow a multi-modal examination including QSM, R2*, volume, diffusion tensor imaging parameters and cerebral blood flow of these two key midbrain nuclei across groups (both PD/control and across cognitive subgroups), and in the evolution of cognitive and motor decline in PD, which we present in detail in the next chapter.

6 Combined Assessment of QSM, R2*, Volume, Diffusion Parameters, and Cerebral Blood Flow in Parkinson's Disease

Magnetic resonance imaging has been widely used in the analysis and investigation of Parkinson's disease and is an active area of research. While the diagnosis of PD remains clinical, advanced MRI methods such as quantitative susceptibility mapping, relaxation rates (R2*(1/T2*)), diffusion tensor imaging and arterial spin labelling may be helpful in providing additional non-invasive and sensitive biomarkers of neurodegeneration in Parkinson's disease.

Quantitative susceptibility mapping measures susceptibility, allowing quantification of brain iron, along with R2*, across the whole brain. Previously in this thesis, higher susceptibility was identified in the SN in a subgroup of PD participants, consistent with other reports of increased iron content in the substantia nigra of PD patients (Barbosa et al., 2015; Du et al., 2016; Murakami et al., 2015). DTI parameters such as fractional anisotropy, mean diffusivity, axial diffusivity, and radial diffusivity help in probing the

underlying microstructural brain changes with progressive disease. Studies have demonstrated decreases in anisotropy in the substantia nigra of PD patients compared to healthy controls (Knossalla et al., 2018; Loane et al., 2016; Péran et al., 2010) and increased diffusivity in the basal ganglia (Loane et al., 2016; Pelizzari et al., 2019; Safai et al., 2020). Similarly, ASL allows quantification of cerebral blood flow, with previous reports of altered perfusion in PD (Pelizzari et al., 2019; Wei et al., 2016).

These different MR parameters, measuring complementary tissue characteristics such as iron deposition, volume, microstructural changes, and perfusion parameters in PD, could be helpful in identifying and describing pathological brain changes in vivo, and ideally, may contribute to an imaging signature capable of monitoring disease progression. Several studies (Aquino et al., 2014; Guangwei Du et al., 2011; Pelizzari et al., 2019; Péran et al., 2010; Safai et al., 2020; Wei et al., 2016) have combined different MR parameters to study midbrain neurodegeneration in Parkinson's disease and reported that amalgamation of MR parameters provides insight to different neuro-pathological process involved in the disease, providing superior information than single MR parameter in isolation.

To the best of our knowledge, no previous PD studies have extended a multi-modal investigation to include QSM, R2*, volume, DTI (FA, AD, MD, and RD) and CBF metrics, but this section of the thesis falls squarely within the priorities of the research community, which encourage multi-modal MRI investigations of specific diseases. Therefore, we simultaneously measured QSM, R2*, volume, DTI and perfusion in substantia nigra and red nucleus in both healthy controls and Parkinson's disease patients. This chapter aims to investigate changes occurring in PD patients compared with healthy controls, and explore any relationship with cognitive and motor function in Parkinson's disease. In addition, this work used multi-modal MRI from the substantia nigra and red nucleus to classify participants as either PD or control; a number of different machine learning algorithms were tested and showed good accuracy. The experiments in this chapter were performed using the database provided by the New Zealand Brain Research Institute consisting a total of 174 participants (47 healthy controls and 127 PD). Of the 127 PD participants, 46 had normal cognition, 66 had mild cognitive impairment, and 15 had dementia.

6.1 Segmentation of Substantia Nigra and Red Nucleus

Substantia nigra and red nucleus segmentations were performed on QSM images using the methodology described in Chapter 5 for all the 174 participants involved in the study. The segmented bilateral SN and RN were mapped to R2*, FA, AD, MD, RD and CBF maps (based on a rigid body coregistration across all images, using nearest-neighbour interpolation). Median values of QSM, R2*, FA, AD, MD, RD and CBF were extracted from left and right SN and RN independently for each individual. Approximate volumes of the SN and RN for each subject were also obtained as the sum of SN and RN areas multiplied by the slice thickness (left and right were computed independently).

6.2 Parkinson's Disease-Related Differences in Susceptibility, Relaxation Rates, Anisotropy, Diffusivity, Perfusion, and Volume

As a first step, this thesis aimed to describe the relationship between multi-modal MRI metrics and measures of PD severity. Median QSM, R2*, FA, AD, MD, RD, CBF and approximate volume from SN and RN ROIs were compared across group (PD/Control), diagnosis (categorical controls/PD-N/PD-MCI/PDD), global cognitive ability, and motor impairment using Bayesian regression models. Bayesian models were fitted using the "brms" (v2.9.0) package in R (v3.6.3). In each model, four chains with 5,000 iterations each were used to generate the posterior sample. Given previous research, we examined hypotheses of (1) increased QSM, R2*, AD, MD, and RD, and (2) decreased volume, FA and CBF in PD. Furthermore, we investigated these metrics across the cognitive subgroups (Controls, PD-N, PD-MCI, and PDD). We specifically tested for a negative association (among QSM, R2*, AD, MD, and RD) with global cognitive ability and a positive association with motor impairment; for volume, FA and CBF, a positive association with cognitive ability and a negative association with motor impairment was tested. Parameter estimates, 95% uncertainty intervals, and probabilities of hypotheses being true are presented.

(1) Group differences: We first tested for evidence of varying susceptibility (QSM value), relaxation rates (R2* value), anisotropy (FA values), diffusivity (AD, MD and RD values), perfusion (CBF values) and volume between PD and healthy

control groups. We modelled each MRI metric independently as a function of age, sex, group (PD/control), side (left or right), ROI (SN/RN), and the interactions age-by-ROI (allowing the relationship between age and MRI metric to vary by ROI), group-by-ROI (allowing the difference in MRI metric between PD and controls to vary between ROI), and group-by-side (allowing the PD effect to vary between left and right side), with an intercept varying by subject, side, and ROI.

(2) Diagnosis: We also tested for evidence of QSM, R2*, FA, AD, MD, RD, CBF and volume differences among healthy controls, PD-N, PD-MCI, and PDD groups. Individual models were used to model each MRI metric as a function of age, sex, diagnosis (PD-N/PD-MCI/PDD/control), side (left or right), ROI (SN/RN), and the interactions age-by-ROI, diagnosis-by-ROI, and diagnosis-by-side, with an intercept varying by subject, side, and ROI.

(3) Cognitive ability: In PD participants only, we modelled MRI metrics as a function of age, sex, side (left or right), ROI (SN/RN), global cognitive ability (aggregate cognitive z score), and the interactions age-by-ROI, cognitive score-by-ROI, and cognitive score-by-side, with an intercept varying by subject, side, and ROI.

(4) Motor impairment: In PD participants only, we modelled MRI metrics as a function of age, sex, side (left or right), ROI (SN/RN), UPDRS III, and the interactions age-by-ROI, UPDRS III-by-ROI, and UPDRS III-by-side, with an intercept varying by subject, side, and ROI.

6.3 Predictive Analysis

Section 6.2 specifically set out to describe the relationship between multiple MRI metrics within the SN and RN and various aspects of PD. However, a different yet related objective of this thesis was to discriminate healthy controls and Parkinson's disease patients using MRI metrics from the SN and RN in a classification/predictive sense. To do so, we employed a number of different predictive/machine learning algorithms. Distinct from previous analyses which aimed to describe how these MRI metrics behave within the SN and RN, here we specifically used the MRI metrics to develop a classification model to determine whether an individual was PD or control based on their MRI from the SN and RN. In this section, we first trained a classification model using a number of different classifiers; we then tested the model using cross validation. To start the process, linear regression was performed to remove the effect of

age and sex for all the modalities; for volume, an additional effect of intracranial volume (ICV) was also removed. In our study, linear support vector machine (Noble, 2006) was used for differentiating healthy controls and PD. SVMs are supervised learning models used for classification in machine learning and are frequently used in neuroimaging (Mateos-Pérez et al., 2018; Orrù, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012).

In this study, leave-one-out cross validation was used for training and testing on a total of 174 subjects. Thirty-two normalized residuals data (obtained after performing linear regression) from multiple modalities including QSM, R2*, FA, AD, MD, RD, CBF, and volume from left and right SN and RN were used as the feature set for the linear SVM. At every step, 10 fold cross validation was performed on the training data with a varying strength of regularization (Churchill, Yourganov, & Strother, 2014; Varoquaux et al., 2017). The final model was selected with the regularization strength that yields minimum validation error. Receiver operating characteristic curve and precision-recall (PR) curve were performed to determine whether the combination of multiple modalities measurements could reliably distinguish PD and healthy controls. Precisely, the area under the receiver operating characteristics curve (AUCROC) and the area under the precision-recall curve (AUCPR) were calculated as a measure of discrimination and to evaluate the model's performance. Moreover, we also trained the model selecting just the five most important features (out of thirty-two) computed using Relief-based algorithms (ReliefF) (Robnik-Šikonja & Kononenko, 2003; Urbanowicz, Meeker, La Cava, Olson, & Moore, 2018) to evaluate the performance. Additionally, non-linear SVM (Gaussian kernel), logistic regression, and random forest (number of trees set to 100) were also used for comparison in this work. All these classifiers were also implemented using the same nested cross-validation approach to separate PD from healthy controls.

6.4 Results

The demographic and clinical data used in this chapter are summarized in Table 6.1.

Table 6.1: Demographic and Clinical Information

Measure	Control (n = 47)	PD-N (n = 46)	PD-MCI (n = 66)	PDD (n = 15)
Sex (M:F)	32:15	30:16	47:19	14:1
Age	74 (8)	69 (8)	73 (6)	76 (6)
Global cognitive ability	0.9 (0.5)	0.4 (0.5)	-0.8 (0.4)	-1.9 (0.5)
MoCA	27 (2)	26 (2)	23 (3)	16 (5)
UPDRS-III	-	34 (14)	41 (14)	51 (9)
Duration of diagnosis (years)	-	9 (4)	8 (5)	8 (4)

Data are mean (Standard Deviation). PD-N: Parkinson's disease with normal cognition; PD-MCI: Parkinson's disease with mild cognitive impairment; PDD: Parkinson's disease dementia. Sex is reported as a ratio. M: Male; F: Female. MoCA: Montreal Cognitive Assessment; UPDRS-III: Unified Parkinson's Disease Rating Scale Part 3. Global cognitive ability was calculated as an aggregate of standardized scores across five cognitive domains.

6.4.1 QSM

Table 6.2 and Figure 6.1 report the results from comparisons between healthy control and PD groups in SN and RN using the proposed segmentation method. We observed strong evidence of increased QSM values in the SN in the PD group relative to controls (probability P = 100% of the PD group having higher QSM values than the control group), with moderate evidence within the RN (P = 98%). Furthermore, we found moderate evidence for an association with age in the SN (probability P = 91% of a positive association), but no evidence in the RN (P = 7%), and weak evidence of a difference between anatomic side (probability P = 85% probability of right side having greater QSM values).

Table 6.2 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed strong evidence of increased QSM values in the SN in the PD-N, PD-MCI and PDD group relative to controls (P = 100%), with moderate evidence within RN (P = 99% (PD-N), P = 90% (PD-MCI), P = 88% (PDD)). Furthermore, we found no evidence of increased QSM values in the SN in PD-MCI and PDD group relative to PD-N (P = 7% (PD-MCI), P = 63% (PDD)) and in the RN (P = 7% (PD-MCI), P = 28% (PDD)). However, there was moderate evidence of increased QSM values in the SN in PDD relative to PD-MCI (P = 91%), but no evidence within RN (P = 66%).

Table 6.2, Figure 6.2 and Figure 6.3 report the relationship between QSM, global cognitive ability, and motor impairment. There was no evidence of a negative relationship between QSM and global cognitive z score in SN (60% probability of a negative association between QSM and global cognitive ability) and within RN (41%). We did not observe evidence for a positive association with motor impairment in SN (68%), but weak evidence in the RN (76%).

Table 6.2: QSM differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between QSM, global cognitive ability, and motor impairment

	QSM	
Region	SN	RN
Control group (ppb)	121 [111, 131]	110 [100, 120]
PD vs controls (ppb)	20 [10, 29]	10 [0.5, 20]
Posterior probability of PD > controls	100%	98%
PD-N vs controls (ppb)	24 [12, 37]	15 [3.2, 27]
Posterior probability of PD-N > controls	100%	99%
PD-MCI vs controls (ppb)	16 [5.1, 27]	6.8 [-3.8, 18]
Posterior probability of PD-MCI > controls	100%	90%
PDD vs controls (ppb)	27 [10, 45]	10 [-7, 27]
Posterior probability of PDD > controls	100%	88%
PD-MCI vs PD-N (ppb)	-8.6 [-20, 2.5]	-8.4 [-20, 2.9]
Posterior probability of PD-MCI > PD-N	7%	7%
PDD vs PD-N (ppb)	3 [-15, 21]	-5.1 [-23, 12]
Posterior probability of PDD > PD-N	63%	28%
PDD vs PD-MCI (ppb)	12 [-5.4, 28]	3.4 [-13, 20]
Posterior probability of PDD > PD-MCI	91%	66%
Global cognitive ability (ppb/point)	-0.8 [-7.7, 6.1]	0.8 [-5.9, 7.6]
Posterior probability of a negative association	60%	41%
UPDRS-III (ppb/point)	0.9 [-2.9, 4.8]	1.4 [-2.5, 5.3]
Posterior probability of a positive association	68%	76%

ppb = parts per billion. For controls, the reported value [95% uncertainty interval] represents the estimate for QSM. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: parts per billion per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

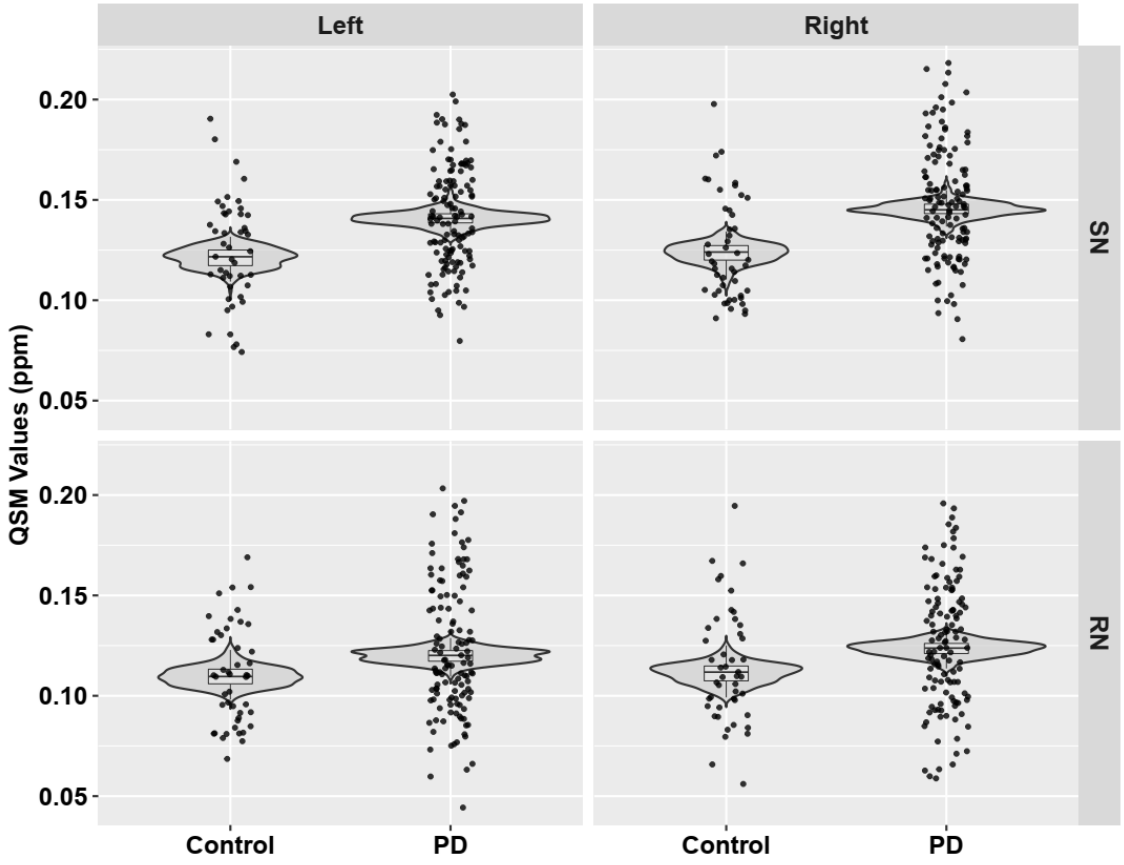


Figure 6.1: QSM values of left and right sides of SN and RN between healthy control and PD. Dots are raw QSM values. Violin plots in gray and boxplots represent the distribution of QSM values adjusted for covariates from 200 draws from the posterior probability distribution.

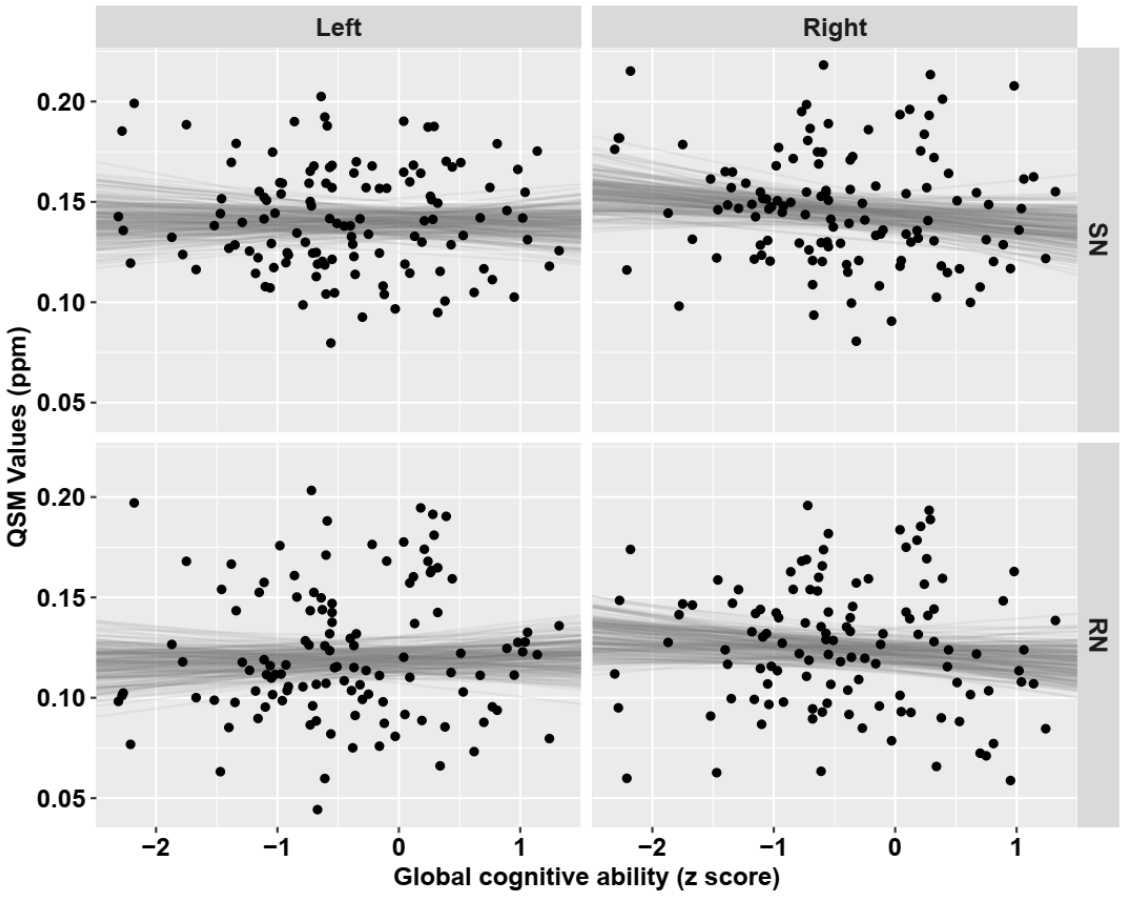


Figure 6.2: Correlations between QSM values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

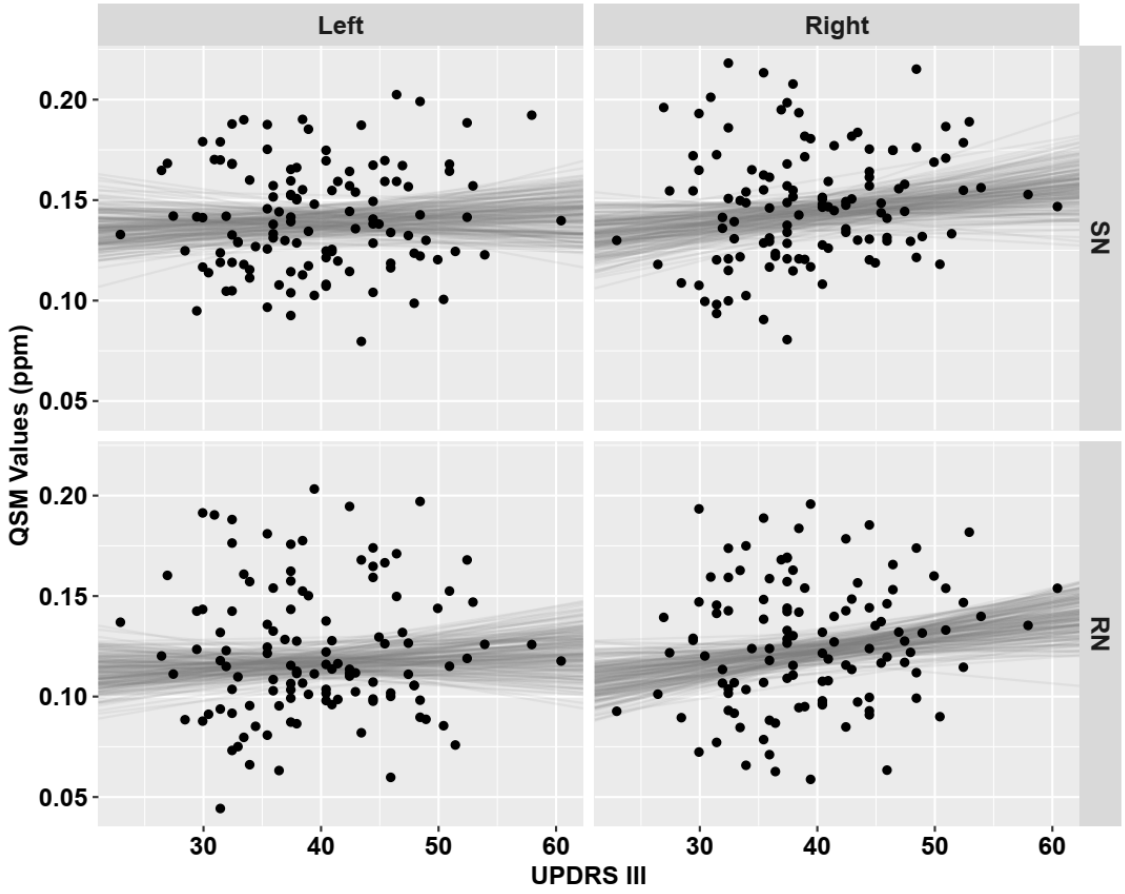


Figure 6.3: Correlations between QSM values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

*6.4.2 R2**

Table 6.3 and Figure 6.4 report the results from comparisons between healthy control and PD groups in SN and RN. We observed strong evidence of increased R2* values in the SN in the PD group relative to controls (probability P = 100%), with moderate evidence within the RN (P = 95%). Furthermore, we found moderate evidence for an association with age in the SN (probability P = 94% of a positive association), with weak evidence in the RN (P = 78%), and no evidence of a difference between anatomic side (probability P = 30% probability of right side having greater R2* values).

Table 6.3 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed strong evidence of increased R2* values in the SN in the PD-N, PD-MCI and PDD group relative to controls (P = 100%). There was also strong evidence of increased R2* values in the RN in the PD-N relative to controls (P = 100%), but no evidence in the PD-MCI (P = 72%), and weak evidence in the PDD (P = 82%). Furthermore, we found no evidence of increased R2* values in the SN in PD-MCI and PDD group relative to PD-N (P = 12% (PD-MCI), P = 72% (PDD)) and in the RN (P = 1% (PD-MCI), P = 15% (PDD)). However, there was moderate evidence of increased R2* values in the SN in PDD relative to PD-MCI (P = 92%), but no evidence within RN (P = 71%).

Table 6.3, Figure 6.5 and Figure 6.6 report the relationship between R2*, global cognitive ability, and motor impairment. We found no evidence of a negative relationship between R2* and global cognitive z score in SN (47% probability of a negative association between R2* and global cognitive ability) and within RN (3%). There was also no evidence for a positive association with motor impairment in SN (71%), and in the RN (63%).

Table 6.3: R2* differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between R2*, global cognitive ability, and motor impairment

	R2*	
Region	SN	RN
Control group (1/sec)	35 [34, 37]	34 [33, 36]
PD vs controls (1/sec)	4 [2.3, 5.7]	1.4 [-0.3, 3.1]
Posterior probability of PD > controls	100%	95%
PD-N vs controls (1/sec)	4.7 [2.5, 6.8]	3 [0.9, 5.1]
Posterior probability of PD-N > controls	100%	100%
PD-MCI vs controls (1/sec)	3.5 [1.6, 5.4]	0.6 [-1.3, 2.4]
Posterior probability of PD-MCI > controls	100%	72%
PDD vs controls (1/sec)	5.6 [2.6, 8.7]	1.4 [-1.5, 4.3]
Posterior probability of PDD > controls	100%	82%
PD-MCI vs PD-N (1/sec)	-1.2 [-3.2, 0.8]	-2.4 [-4.4, -0.5]
Posterior probability of PD-MCI > PD-N	12%	1%
PDD vs PD-N (1/sec)	0.9 [-2.2, 4]	-1.6 [-4.8, 1.5]
Posterior probability of PDD > PD-N	72%	15%
PDD vs PD-MCI (1/sec)	2.1 [-0.8, 5]	0.8 [-2.1, 3.7]
Posterior probability of PDD > PD-MCI	92%	71%
Global cognitive ability (1/sec/point)	0.05 [-1.2, 1.2]	1.1 [-0.03, 2.3]
Posterior probability of a negative association	47%	3%
UPDRS-III (1/sec/point)	0.4 [-0.5, 0.9]	0.1 [-0.6, 0.8]
Posterior probability of a positive association	71%	63%

For controls, the reported value [95% uncertainty interval] represents the estimate for R2*. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: 1 per sec per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

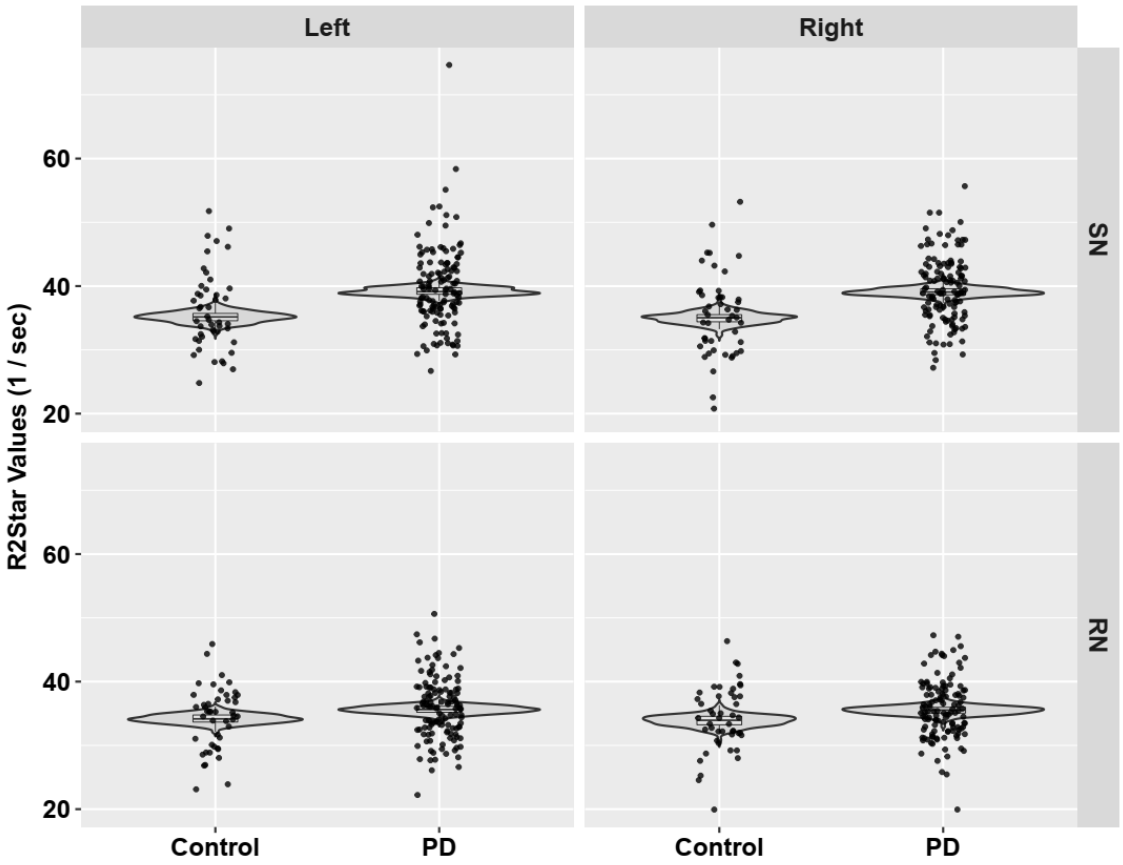


Figure 6.4: R2* values of left and right sides of SN and RN between healthy control and PD. Dots are raw R2* values. Violin plots in gray and boxplots represent the distribution of R2* values adjusted for covariates from 200 draws from the posterior probability distribution.

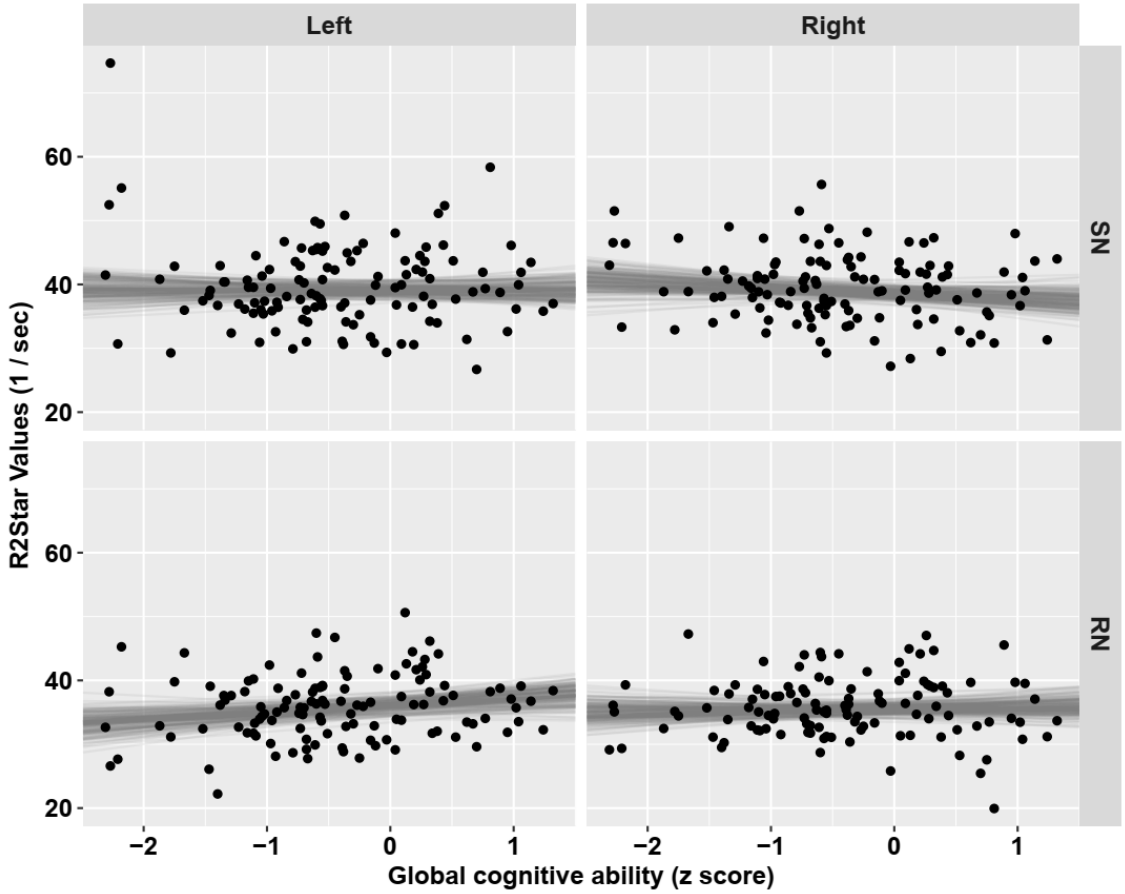


Figure 6.5: Correlations between R2* values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

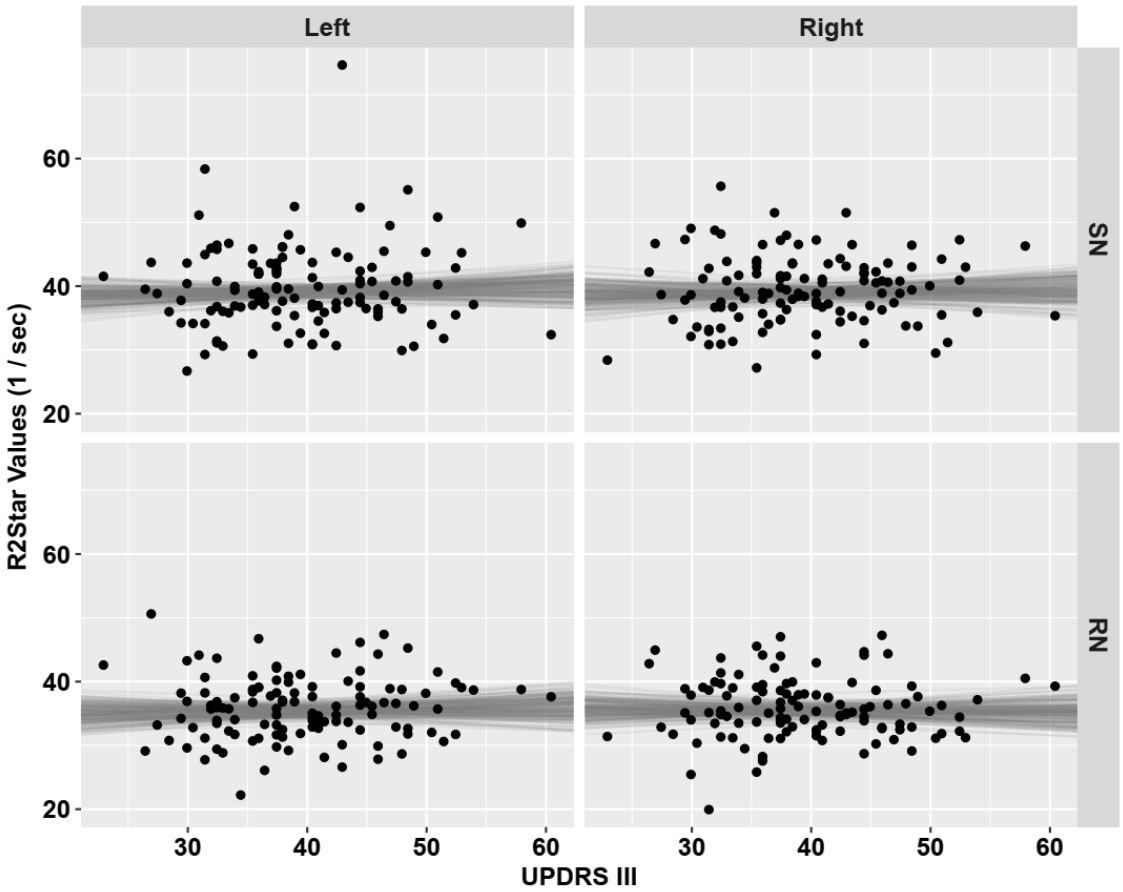


Figure 6.6: Correlations between R2* values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

6.4.3 Volume

Table 6.4 and Figure 6.7 report the results from comparisons between healthy control and PD groups in SN and RN. We observed no evidence of decreased volume in the SN in the PD group relative to controls (probability P = 1%), indicating *increased* volume in the SN in PD, and within the RN (P = 65%); this slightly counter-intuitive finding is discussed in Section 6.5. Furthermore, we found no evidence for an association with age in the SN (probability P = 63% of a negative association), but weak evidence in the RN (P = 77%).

Table 6.4 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed no evidence of decreased volume in the SN in the PD-N (P = 0%), PD-MCI (P = 12%), and PDD (P = 3%) relative to controls. There was no evidence of decreased volume in the RN in the PD-N relative to controls (P = 53%), but weak evidence in the PD-MCI (P = 78%), and no evidence in the PDD (P = 30%). Furthermore, we found moderate evidence of decreased volume in the SN in PD-MCI relative to PD-N (P = 99%), and no evidence in PDD (P = 63%), weak evidence within RN in PD-MCI (P = 75%), and no evidence in PDD (P = 29%). We also did not observe evidence of decreased volume in the SN in PDD relative to PD-MCI (P = 11%) and in the RN (P = 14%).

Table 6.4, Figure 6.8 and Figure 6.9 report the relationship between volume, global cognitive ability, and motor impairment. We found no evidence of a positive relationship between volume and global cognitive z score in SN (62% probability of a positive association between volume and global cognitive ability) and within RN (51%). There was also no evidence for a negative association with motor impairment in SN (18%), and in the RN (19%).

Table 6.4: Volume differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between volume, global cognitive ability, and motor impairment

	Volume	
Region	SN	RN
Control group (mm^3)	1594 [1486, 1701]	551 [454, 647]
PD vs controls (mm^3)	139 [28, 250]	-18 [-112, 76]
Posterior probability of PD < controls	1%	65%
PD-N vs controls (mm^3)	221 [79, 362]	-4.2 [-128, 118]
Posterior probability of PD-N < controls	0%	53%
PD-MCI vs controls (mm^3)	74 [-50, 198]	-42 [-150, 65]
Posterior probability of PD-MCI < controls	12%	78%
PDD vs controls (mm^3)	188 [-11, 384]	44 [-123, 212]
Posterior probability of PDD < controls	3%	30%
PD-MCI vs PD-N (mm^3)	-147 [-265, -23]	-38 [-145, 70]
Posterior probability of PD-MCI < PD-N	99%	75%
PDD vs PD-N (mm^3)	-32 [-229, 169]	48 [-122, 217]
Posterior probability of PDD < PD-N	63%	29%
PDD vs PD-MCI (mm^3)	115 [-72, 299]	86 [-69, 243]
Posterior probability of PDD < PD-MCI	11%	14%
Global cognitive ability (mm^3/point)	11 [-62, 82]	0.7 [-63, 66]
Posterior probability of a positive association	62%	51%
UPDRS-III (mm^3/point)	19 [-21, 59]	16 [-20, 52]
Posterior probability of a negative association	18%	19%

For controls, the reported value [95% uncertainty interval] represents the estimate for volume. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: mm^3 per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

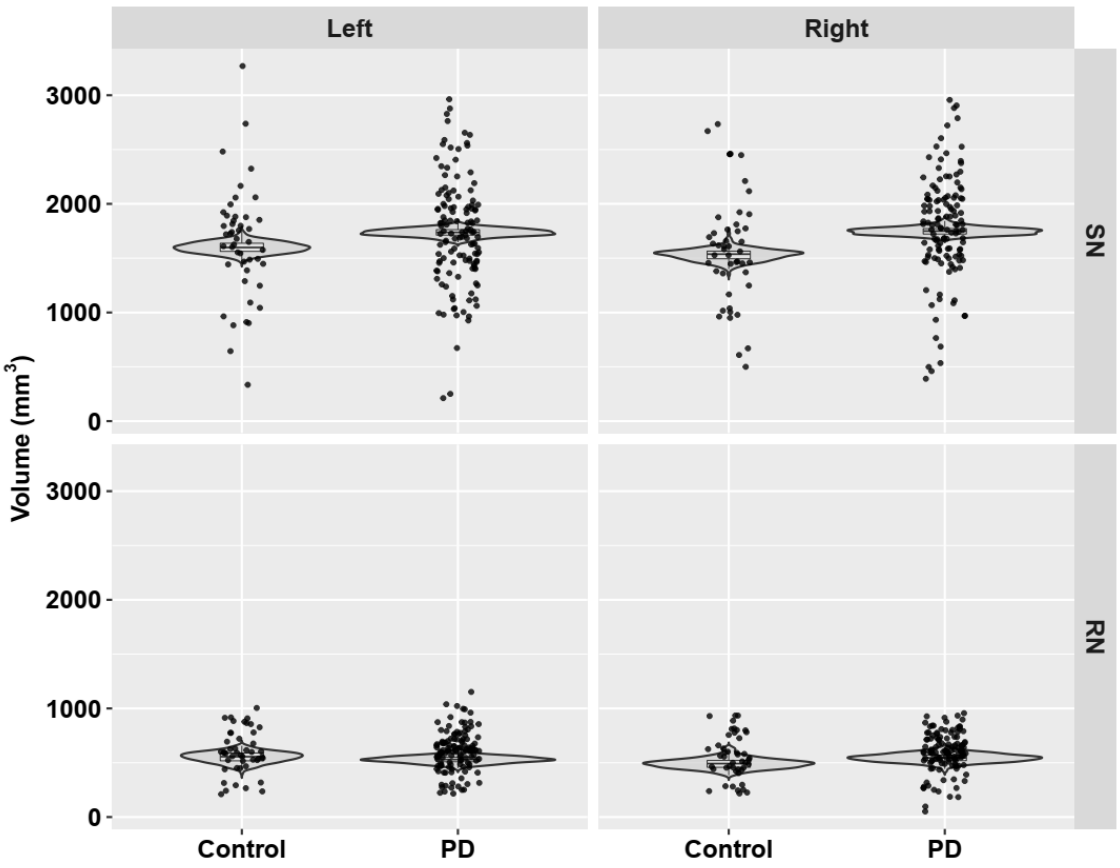


Figure 6.7: Volume of left and right sides of SN and RN between healthy control and PD. Dots are raw volume. Violin plots in gray and boxplots represent the distribution of volume adjusted for covariates from 200 draws from the posterior probability distribution.

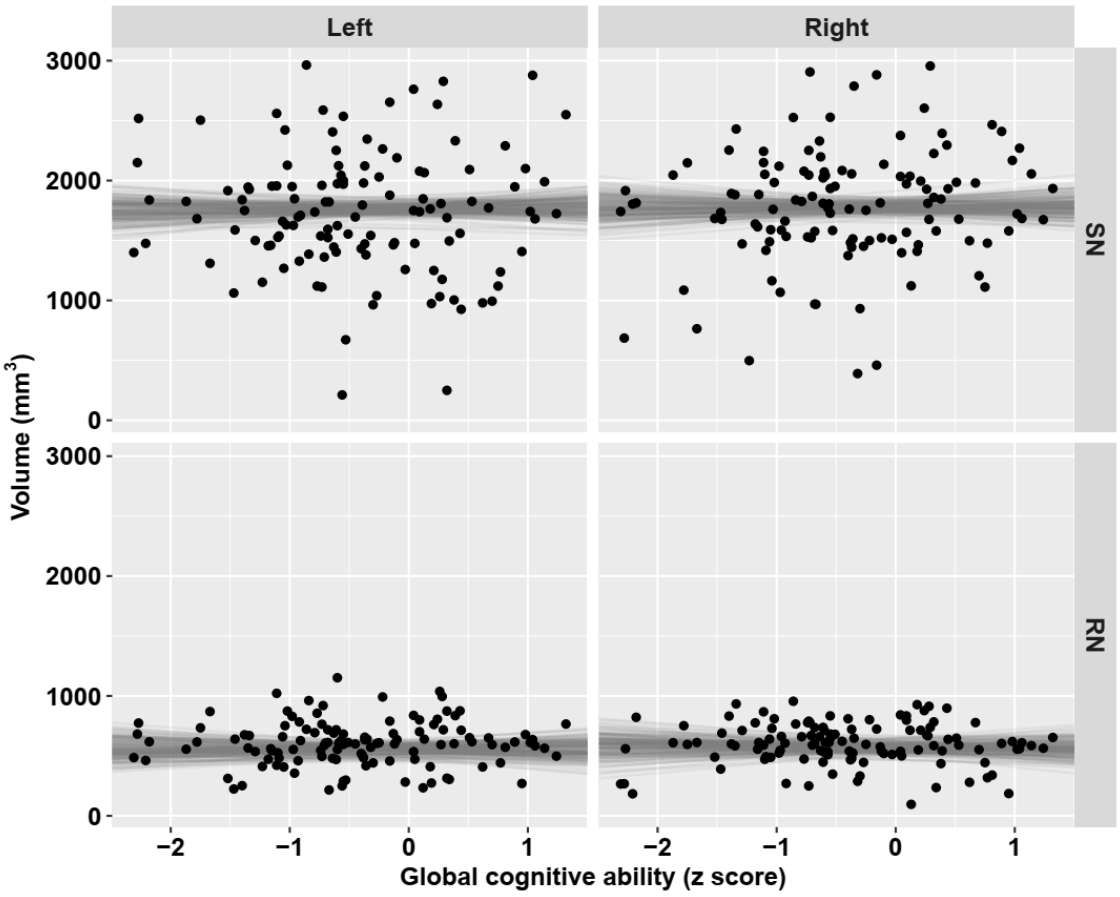


Figure 6.8: Correlations between volume of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

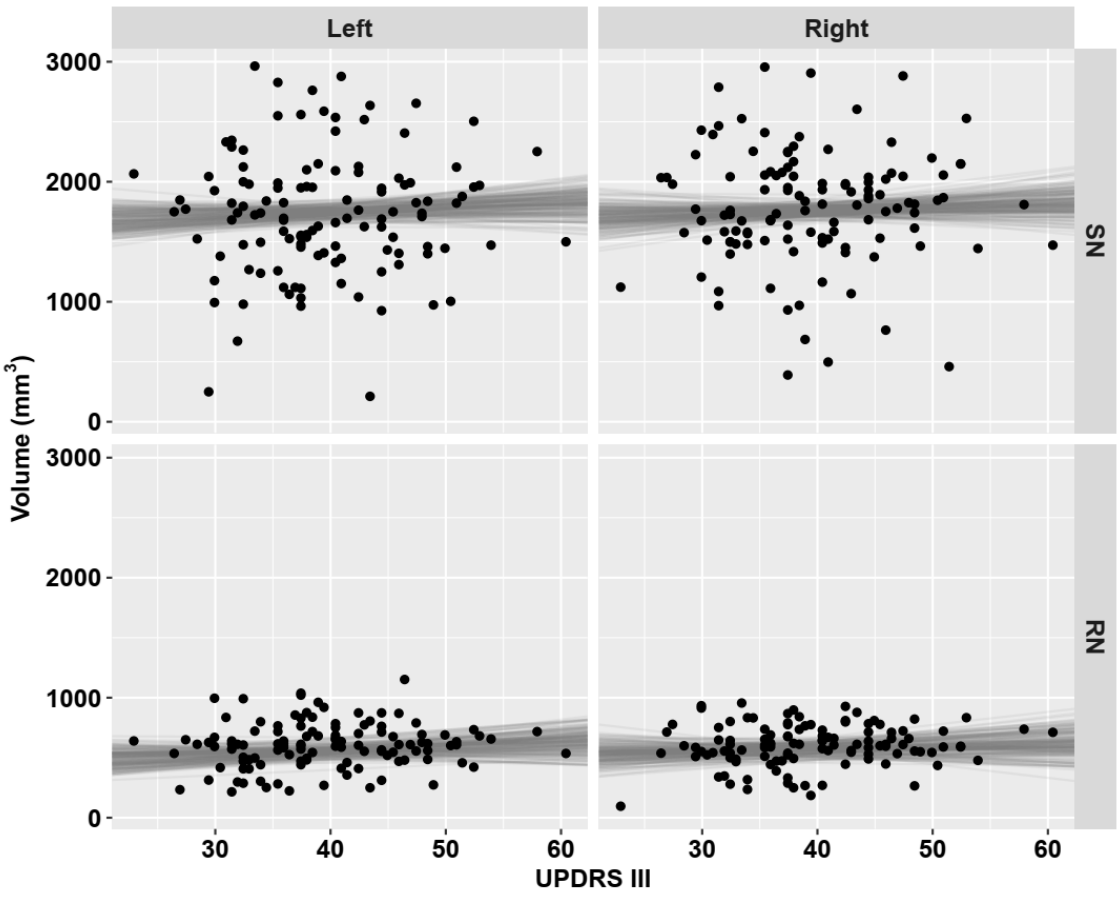


Figure 6.9: Correlations between volume of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

6.4.4 FA

Table 6.5 and Figure 6.10 report the results from comparisons between healthy control and PD groups in SN and RN. We observed moderate evidence of decreased FA values in the SN in the PD group relative to controls (probability P = 88%), and no evidence within the RN (P = 30%). Furthermore, we found no evidence for an association with age in the SN (probability P = 9% of a negative association), and in the RN (P = 24%).

Table 6.5 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed moderate evidence of decreased FA values in the SN in the PD-N relative to controls (P = 97%), weak evidence in PD-MCI (P = 78%), and no evidence in PDD (P = 30%) but there was no evidence of decreased FA values in the RN in PD-N (P = 66%), PD-MCI (P = 26%) and PDD (P = 24%) respectively. Furthermore, we found no evidence of decreased FA values in the SN in PD-MCI (P = 10%) and PDD (P = 3%) relative to PD-N, and within RN in PD-MCI (P = 14%), and in PDD (P = 16%). We also did not observe evidence of decreased FA values in the SN in PDD relative to PD-MCI (P = 14%) and in the RN (P = 38%).

Table 6.5, Figure 6.11 and Figure 6.12 report the relationship between FA, global cognitive ability, and motor impairment. We found no evidence of a positive relationship between FA and global cognitive z score in SN (1% probability of a positive association between FA and global cognitive ability) and within RN (6%). There was also no evidence for a negative association with motor impairment in SN (30%), and in the RN (12%).

Table 6.5: FA differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between FA, global cognitive ability, and motor impairment

	FA	
Region	SN	RN
Control group	400 [381, 419]	383 [363, 404]
PD vs controls	-11 [-31, 7.3]	5.6 [-15, 26]
Posterior probability of PD < controls	88%	30%
PD-N vs controls	-25 [-50, 0.9]	-5.9 [-33, 21]
Posterior probability of PD-N < controls	97%	66%
PD-MCI vs controls	-8.6 [-31, 13]	7.7 [-15, 31]
Posterior probability of PD-MCI < controls	78%	26%
PDD vs controls	9.7 [-25, 45]	13 [-23, 48]
Posterior probability of PDD < controls	30%	24%
PD-MCI vs PD-N	16 [-7.8, 40]	14 [-11, 39]
Posterior probability of PD-MCI < PD-N	10%	14%
PDD vs PD-N	34 [-2, 71]	19 [-18, 55]
Posterior probability of PDD < PD-N	3%	16%
PDD vs PD-MCI	18 [-16, 53]	5.1 [-29, 39]
Posterior probability of PDD < PD-MCI	14%	38%
Global cognitive ability	-17 [-32, -2.4]	-12 [-26, 3.1]
Posterior probability of a positive association	1%	6%
UPDRS-III	2.1 [-5.6, 9.8]	4.8 [-3.1, 13]
Posterior probability of a negative association	30%	12%

For controls, the reported value [95% uncertainty interval] represents the estimate for FA. The values are multiplied by 10^3. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

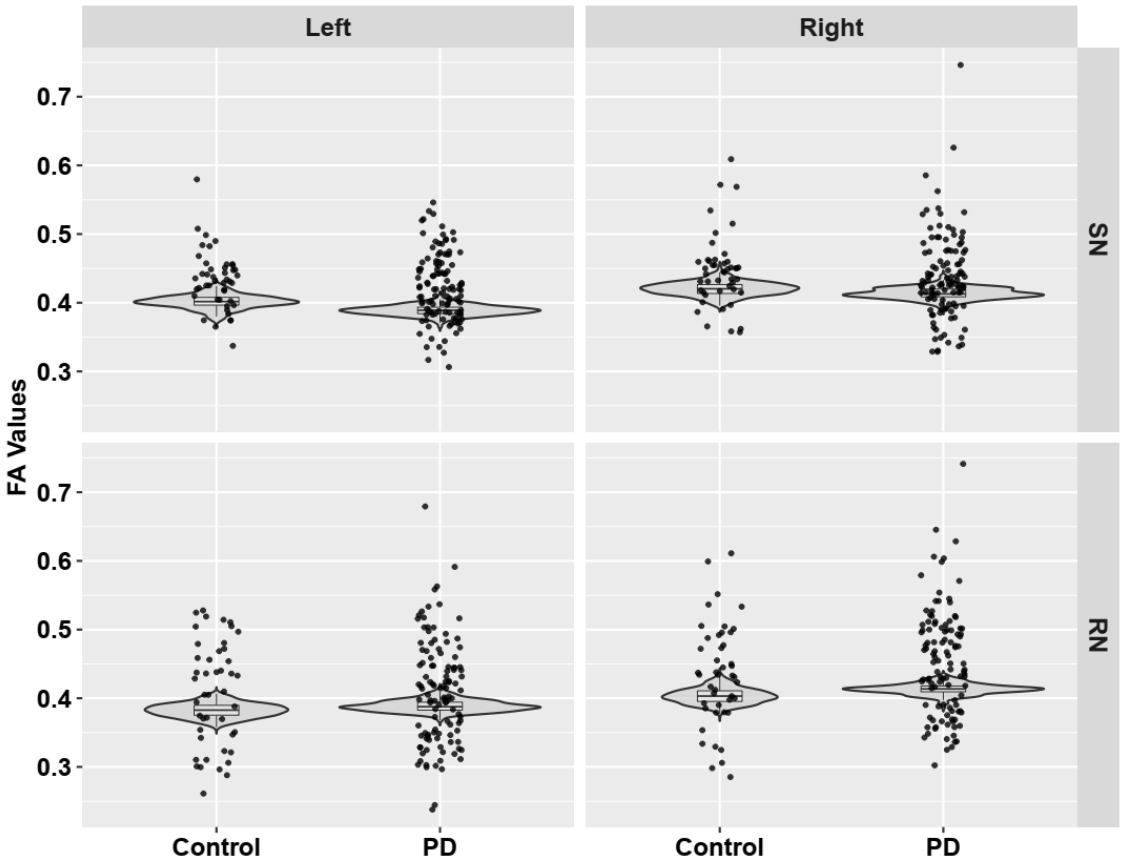


Figure 6.10: FA values of left and right sides of SN and RN between healthy control and PD. Dots are raw FA values. Violin plots in gray and boxplots represent the distribution of FA values adjusted for covariates from 200 draws from the posterior probability distribution.

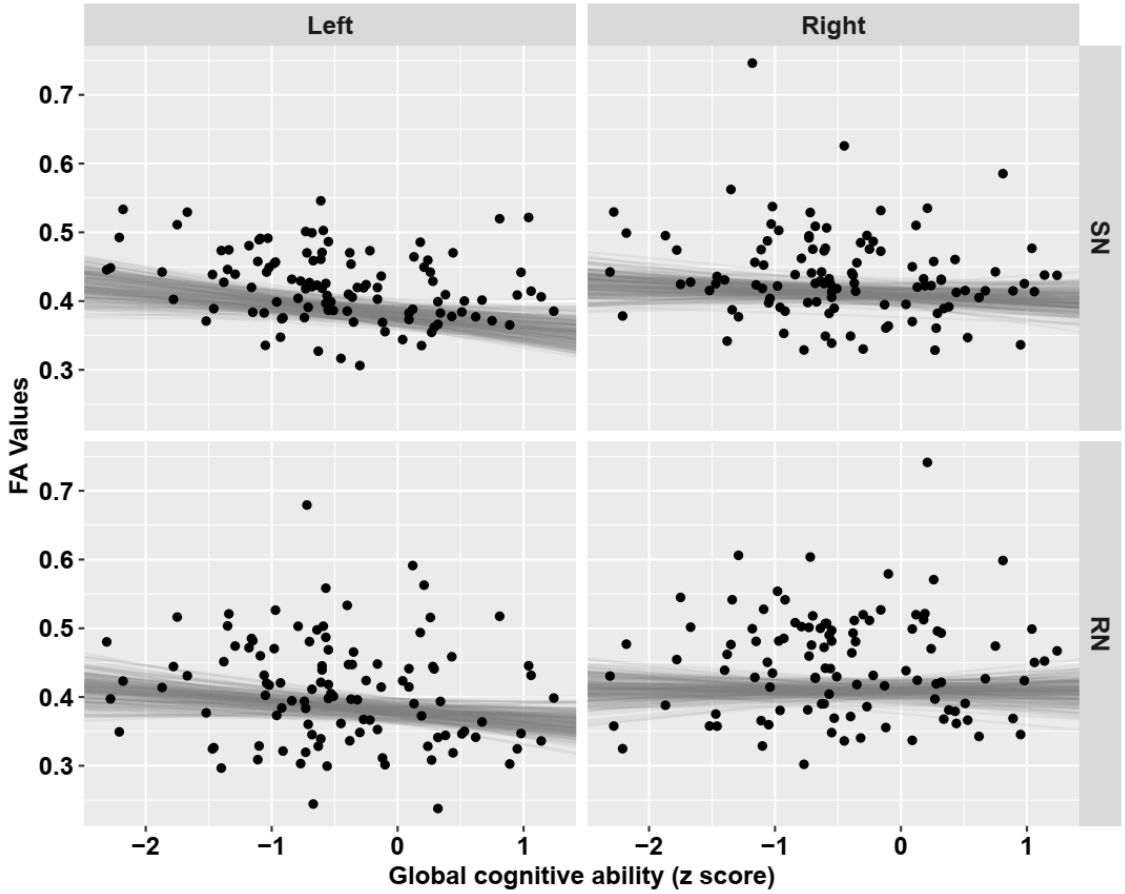


Figure 6.11: Correlations between FA values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

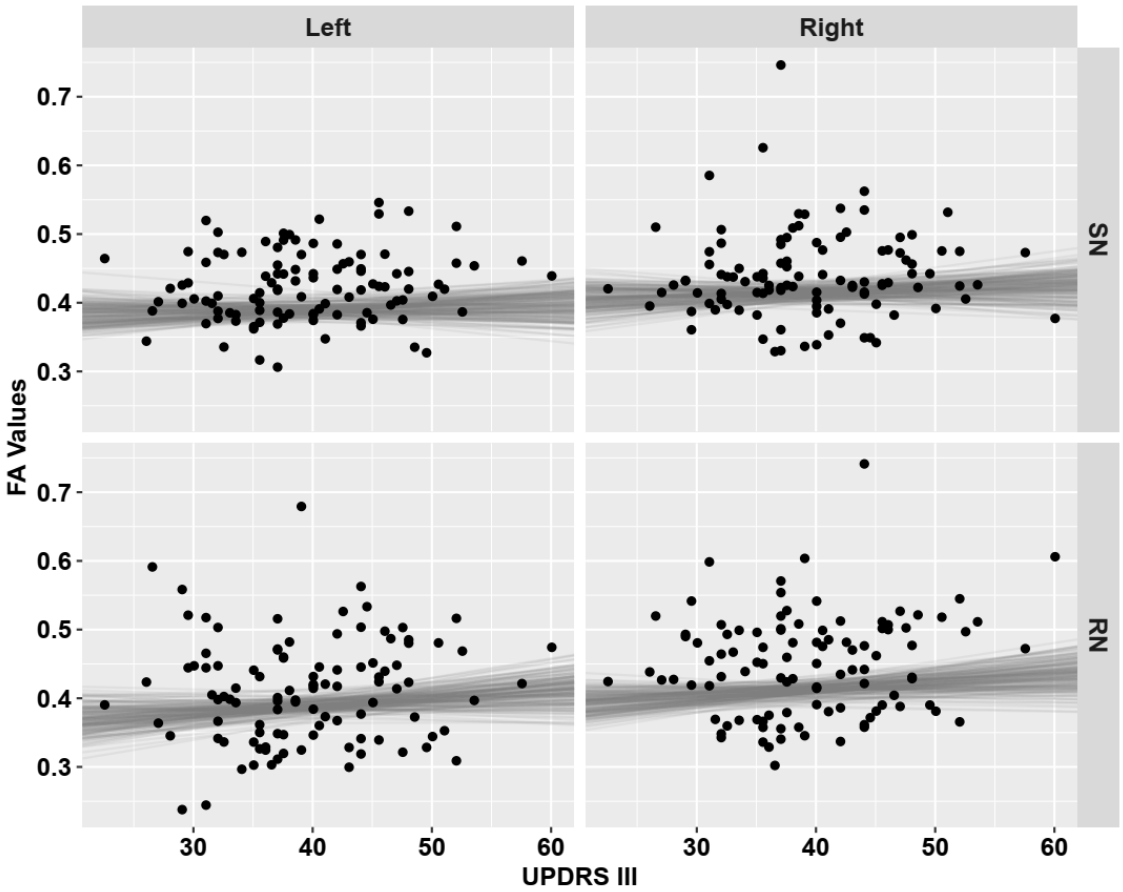


Figure 6.12: Correlations between FA values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

6.4.5 AD

Table 6.6 and Figure 6.13 report the results from comparisons between healthy control and PD groups in SN and RN. We observed moderate evidence of increased AD values in the SN in the PD group relative to controls (probability P = 99%), and no evidence within the RN (P = 71%). Furthermore, we found strong evidence for an association with age in the SN (probability P = 100% of a positive association), with moderate evidence in the RN (P = 91%).

Table 6.6 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed moderate evidence of increased AD values in the SN in the PD-N, PD-MCI and PDD group relative to controls (P = 99% (PD-N), P = 94% (PD-MCI), P = 93% (PDD)). There was no evidence of increased AD values in the RN in the PD-N relative to controls (P = 71%) but weak evidence in the PD-MCI (P = 77%), and no evidence in the PDD (P = 48%). Furthermore, we found no evidence of increased AD values in the SN in PD-MCI and PDD group relative to PD-N (P = 15% (PD-MCI), P = 43% (PDD)), and in the RN (P = 55% (PD-MCI), P = 34% (PDD)). We also observed no evidence of increased AD values in the SN in PDD relative to PD-MCI (P = 70%), and within RN (P = 29%).

Table 6.6, Figure 6.14 and Figure 6.15 report the relationship between AD, global cognitive ability, and motor impairment. We found weak evidence of a negative relationship between AD and global cognitive z score in SN (77% probability of a negative association between AD and global cognitive ability) and within RN (75%). We also observed weak evidence for a positive association with motor impairment in SN (80%), and no evidence in the RN (38%).

Table 6.6: AD differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between AD, global cognitive ability, and motor impairment

	AD	
Region	SN	RN
Control group (mm^2/sec)	129 [124, 133]	119 [115, 124]
PD vs controls (mm^2/sec)	4.9 [0.6, 9.2]	1.2 [-3.2, 5.5]
Posterior probability of PD > controls	99%	71%
PD-N vs controls (mm^2/sec)	6.6 [1.1, 12]	1.6 [-4, 7.2]
Posterior probability of PD-N > controls	99%	71%
PD-MCI vs controls (mm^2/sec)	3.9 [-1.1, 8.7]	1.9 [-3.1, 6.6]
Posterior probability of PD-MCI > controls	94%	77%
PDD vs controls (mm^2/sec)	5.9 [-1.8, 14]	-0.2 [-7.9, 7.5]
Posterior probability of PDD > controls	93%	48%
PD-MCI vs PD-N (mm^2/sec)	-2.8 [-7.9, 2.5]	0.3 [-5.2, 5.6]
Posterior probability of PD-MCI > PD-N	15%	55%
PDD vs PD-N (mm^2/sec)	-0.7 [-8.8, 7.4]	-1.7 [-10, 6.4]
Posterior probability of PDD > PD-N	43%	34%
PDD vs PD-MCI (mm^2/sec)	2.1 [-5.3, 9.7]	-2.1 [-9.5, 5.3]
Posterior probability of PDD > PD-MCI	70%	29%
Global cognitive ability (mm^2/sec/point)	-1.3 [-4.7, 2.1]	-1.2 [-4.5, 2.2]
Posterior probability of a negative association	77%	75%
UPDRS-III (mm^2/sec/point)	0.8 [-1.1, 2.6]	-0.3 [-2.1, 1.5]
Posterior probability of a positive association	80%	38%

For controls, the reported value [95% uncertainty interval] represents the estimate for AD. The values are multiplied by 10^5. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: mm^2 per sec per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

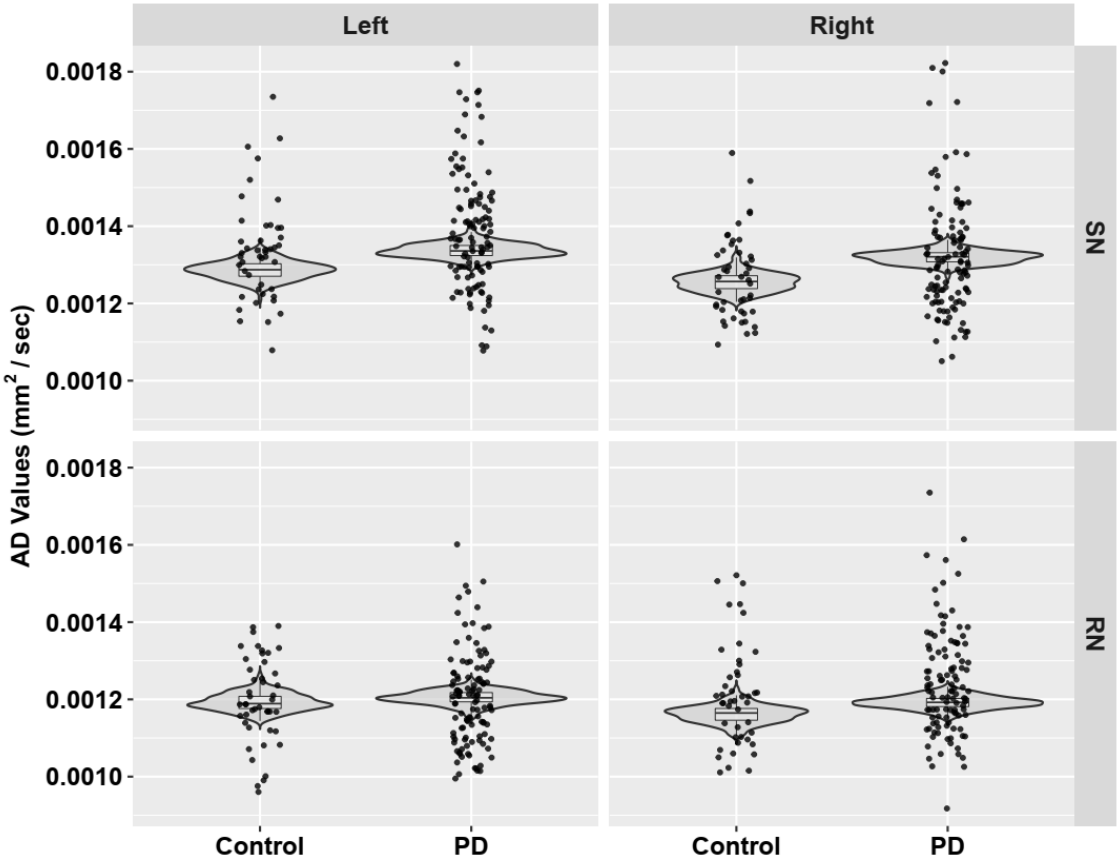


Figure 6.13: AD values of left and right sides of SN and RN between healthy control and PD. Dots are raw AD values. Violin plots in gray and boxplots represent the distribution of AD values adjusted for covariates from 200 draws from the posterior probability distribution.

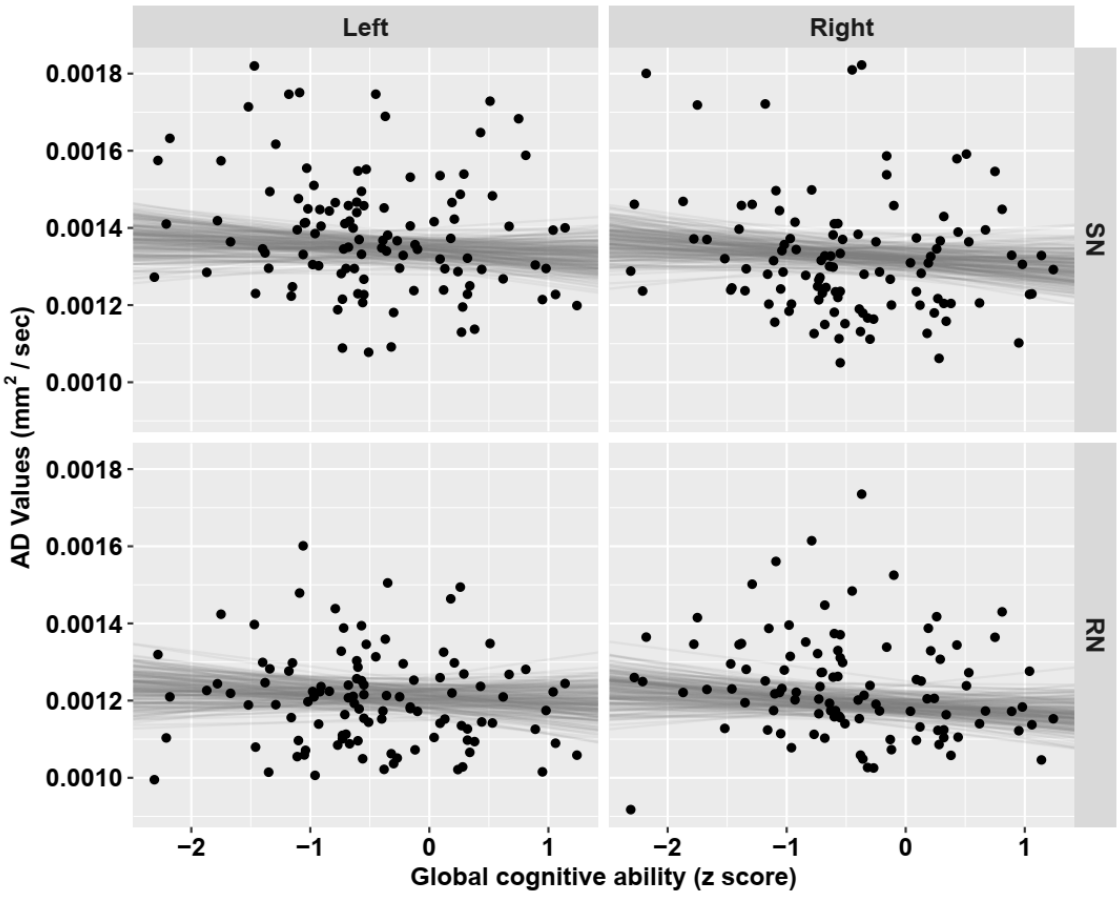


Figure 6.14: Correlations between AD values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

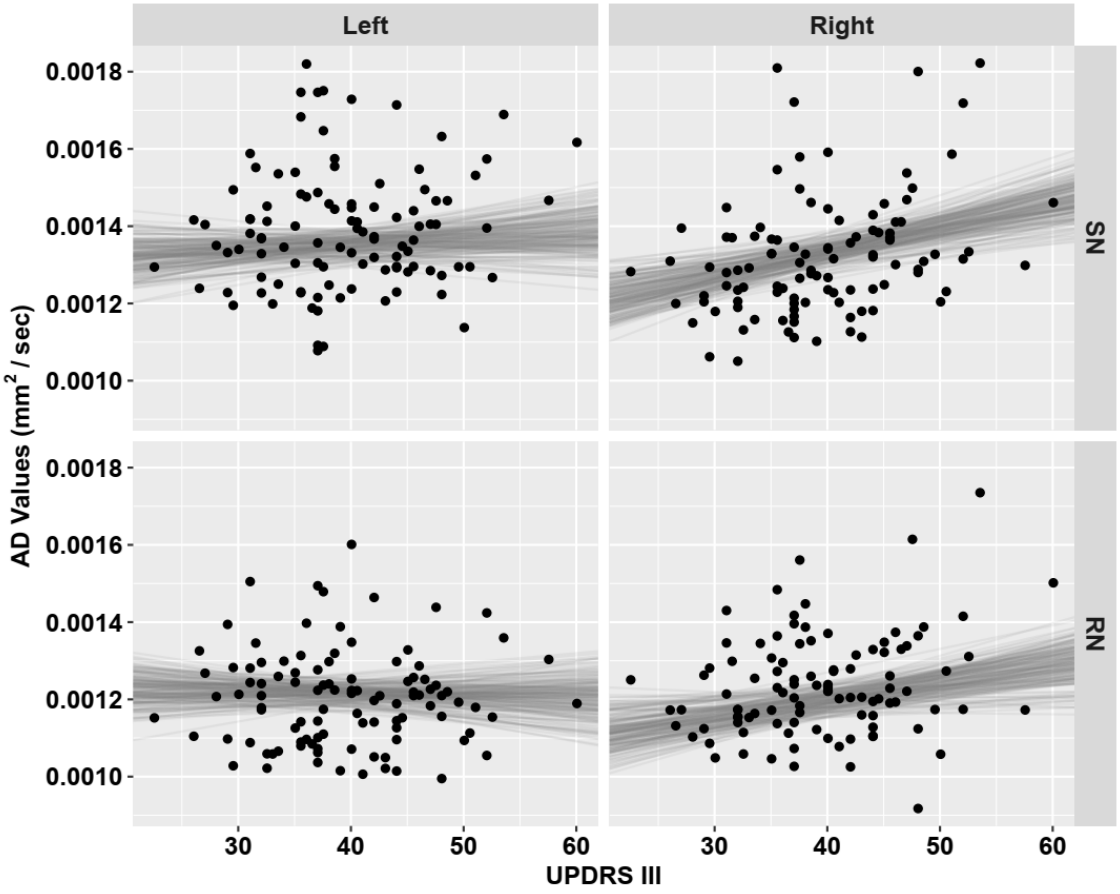


Figure 6.15: Correlations between AD values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

6.4.6 MD

Table 6.7 and Figure 6.16 report the results from comparisons between healthy control and PD groups in SN and RN. We observed strong evidence of increased MD values in the SN in the PD group relative to controls (probability P = 100%), and moderate evidence within the RN (P = 88%). Furthermore, we found moderate evidence for an association with age in the SN (probability P = 99% of a positive association), and weak evidence in the RN (P = 85%).

Table 6.7 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed strong evidence of increased MD values in the SN in PD-N relative to controls (P = 100%), moderate evidence in PD-MCI (P = 97%), and in PDD (P = 94%) respectively. There was moderate evidence of increased MD values in the RN in the PD-N (P = 94%), but weak evidence in the PD-MCI (P = 84%), and no evidence in the PDD (P = 67%) relative to controls. Furthermore, we found no evidence of increased MD values in the SN in PD-MCI and PDD group relative to PD-N (P = 1% (PD-MCI), P = 13% (PDD)) and in the RN (P = 23% (PD-MCI), P = 25% (PDD)). We also observed no evidence of increased MD values in the SN in PDD relative to PD-MCI (P = 64%), and within RN (P = 42%).

Table 6.7, Figure 6.17 and Figure 6.18 report the relationship between MD, global cognitive ability, and motor impairment. We found no evidence of a negative relationship between MD and global cognitive z score in SN (42% probability of a negative association between MD and global cognitive ability) and within RN (64%). However, we observed moderate evidence for a positive association with motor impairment in SN (88%), but no evidence in the RN (45%).

Table 6.7: MD differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between MD, global cognitive ability, and motor impairment

	MD	
Region	SN	RN
Control group (mm^2/sec)	88 [86, 91]	83 [81, 86]
PD vs controls (mm^2/sec)	4 [1.5, 6.6]	1.5 [-1, 4]
Posterior probability of PD > controls	100%	88%
PD-N vs controls (mm^2/sec)	6.3 [3.1, 9.6]	2.6 [-0.7, 5.7]
Posterior probability of PD-N > controls	100%	94%
PD-MCI vs controls (mm^2/sec)	2.8 [0, 5.6]	1.4 [-1.5, 4.2]
Posterior probability of PD-MCI > controls	97%	84%
PDD vs controls (mm^2/sec)	3.6 [-0.8, 7.9]	0.9 [-3.5, 5.3]
Posterior probability of PDD > controls	94%	67%
PD-MCI vs PD-N (mm^2/sec)	-3.5 [-6.5, -0.5]	-1.2 [-4.2, 1.9]
Posterior probability of PD-MCI > PD-N	1%	23%
PDD vs PD-N (mm^2/sec)	-2.7 [-7.4, 2]	-1.6 [-6.2, 3]
Posterior probability of PDD > PD-N	13%	25%
PDD vs PD-MCI (mm^2/sec)	0.8 [-3.5, 5]	-0.5 [-4.6, 3.8]
Posterior probability of PDD > PD-MCI	64%	42%
Global cognitive ability (mm^2/sec/point)	0.2 [-1.8, 2.2]	-0.4 [-2.4, 1.6]
Posterior probability of a negative association	42%	64%
UPDRS-III (mm^2/sec/point)	0.6 [-0.4, 1.6]	-0.1 [-1.1, 1]
Posterior probability of a positive association	88%	45%

For controls, the reported value [95% uncertainty interval] represents the estimate for MD. The values are multiplied by 10^5. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: mm^2 per sec per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

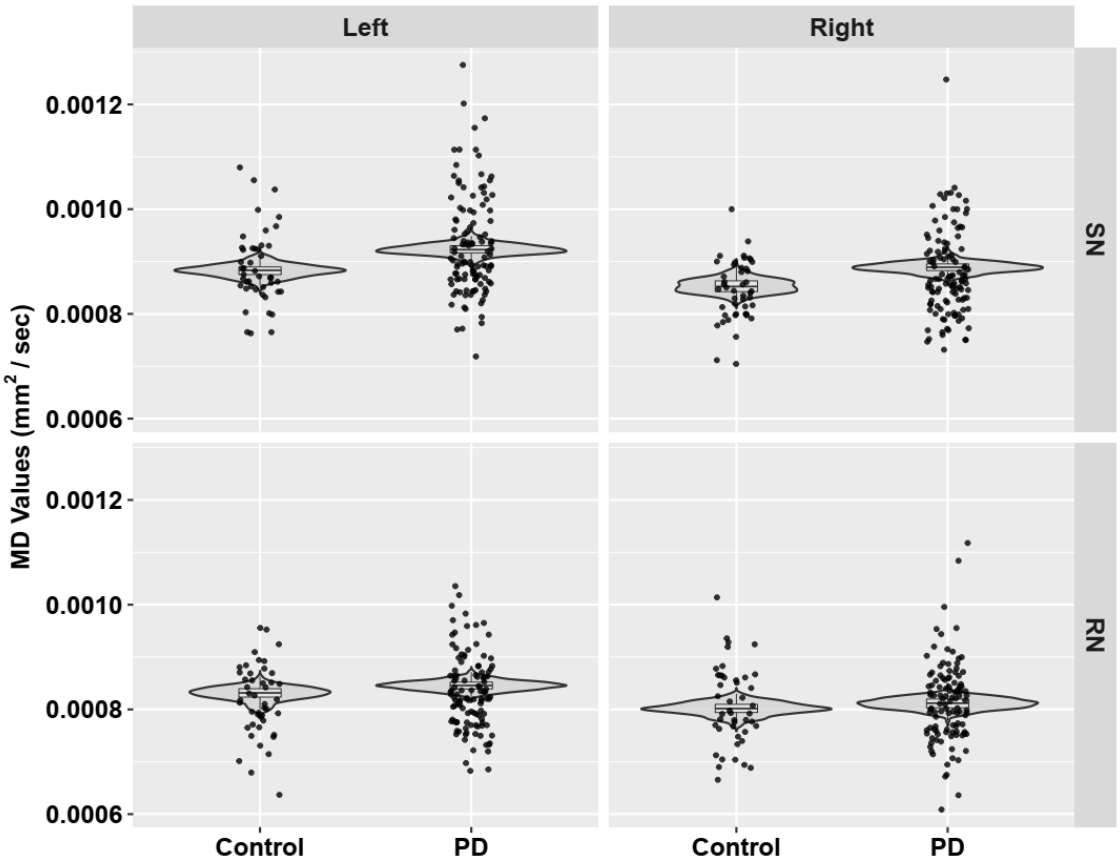


Figure 6.16: MD values of left and right sides of SN and RN between healthy control and PD. Dots are raw MD values. Violin plots in gray and boxplots represent the distribution of MD values adjusted for covariates from 200 draws from the posterior probability distribution.

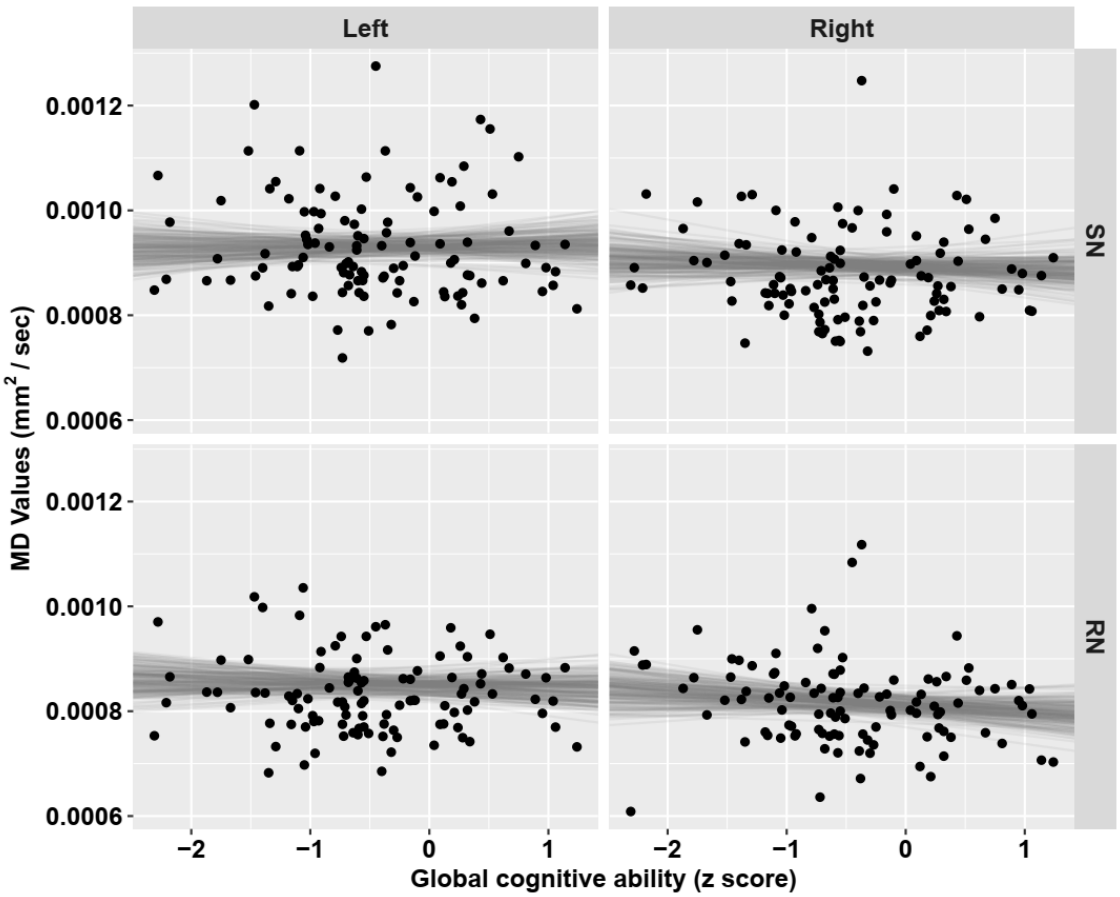


Figure 6.17: Correlations between MD values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

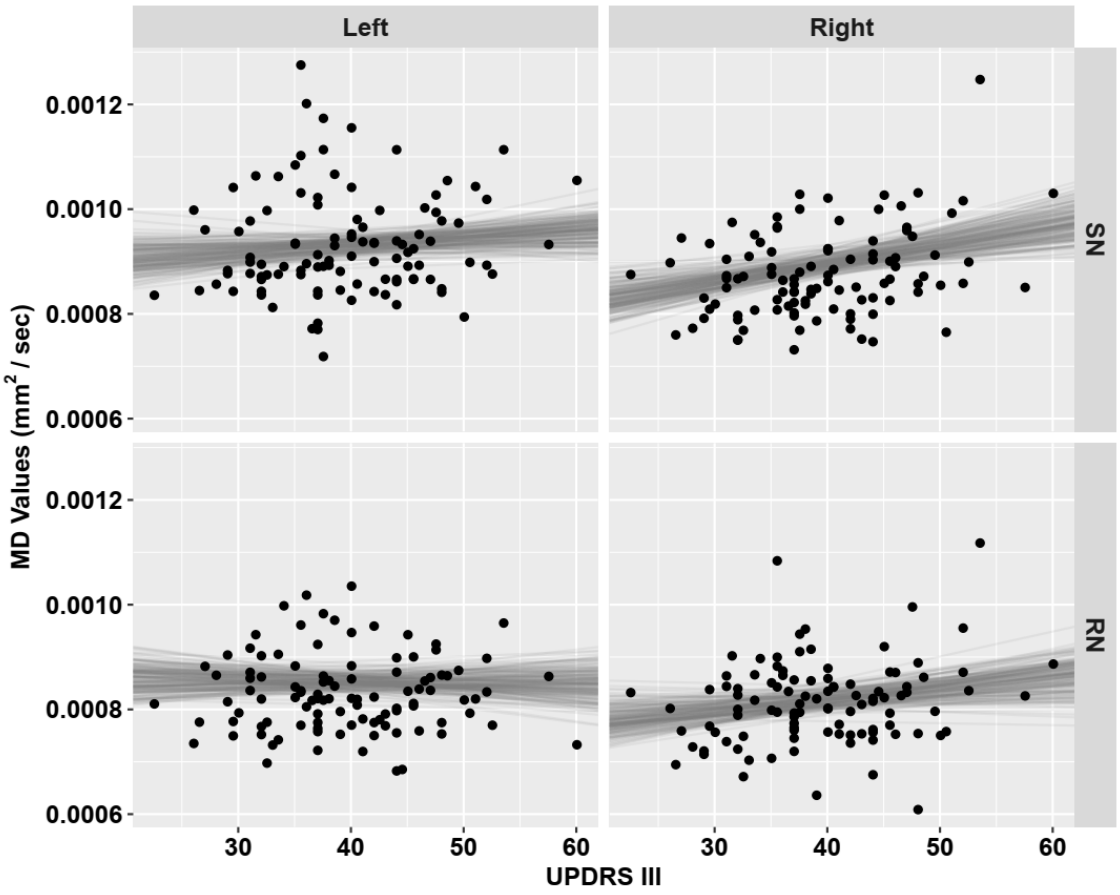


Figure 6.18: Correlations between MD values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

6.4.7 RD

Table 6.8 and Figure 6.19 report the results from comparisons between healthy control and PD groups in SN and RN. We observed strong evidence of increased RD values in the SN in the PD group relative to controls (probability P = 100%), and weak evidence within the RN (P = 78%). Furthermore, we found weak evidence for an association with age in the SN (probability P = 85% of a positive association), and in the RN (P = 76%).

Table 6.8 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed strong evidence of increased RD values in the SN in PD-N relative to controls (P = 100%), moderate evidence in PD-MCI (P = 96%), and weak evidence in PDD (P = 78%). There was moderate evidence of increased RD values in the RN in the PD-N (P = 87%), but no evidence in the PD-MCI (P = 71%), and PDD (P = 59%) relative to controls. Furthermore, we found no evidence of increased RD values in the SN in PD-MCI and PDD group relative to PD-N (P = 2% (PD-MCI), P = 5% (PDD)) and in the RN (P = 25% (PD-MCI), P = 29% (PDD)). We also observed no evidence of increased RD values in the SN in PDD relative to PD-MCI (P = 37%), and within RN (P = 45%).

Table 6.8, Figure 6.20 and Figure 6.21 report the relationship between RD, global cognitive ability, and motor impairment. We found no evidence of a negative relationship between RD and global cognitive z score in SN (11% probability of a negative association between RD and global cognitive ability) and within RN (56%). However, we observed weak evidence for a positive association with motor impairment in SN (79%), but no evidence in the RN (35%).

Table 6.8: RD differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between RD, global cognitive ability, and motor impairment

	RD	
Region	SN	RN
Control group (mm^2/sec)	68 [66, 71]	65 [62, 67]
PD vs controls (mm^2/sec)	3.6 [1.1, 6.1]	1 [-1.5, 3.5]
Posterior probability of PD > controls	100%	78%
PD-N vs controls (mm^2/sec)	5.8 [2.6, 9.1]	1.9 [-1.4, 5.1]
Posterior probability of PD-N > controls	100%	87%
PD-MCI vs controls (mm^2/sec)	2.5 [-0.3, 5.2]	0.8 [-2, 3.7]
Posterior probability of PD-MCI > controls	96%	71%
PDD vs controls (mm^2/sec)	1.8 [-2.8, 6.2]	0.5 [-3.9, 4.9]
Posterior probability of PDD > controls	78%	59%
PD-MCI vs PD-N (mm^2/sec)	-3.3 [-6.4, -0.2]	-1.1 [-4.1, 2]
Posterior probability of PD-MCI > PD-N	2%	25%
PDD vs PD-N (mm^2/sec)	-4.1 [-8.8, 0.6]	-1.3 [-5.9, 3.2]
Posterior probability of PDD > PD-N	5%	29%
PDD vs PD-MCI (mm^2/sec)	-0.8 [-5.1, 3.5]	-0.3 [-4.5, 4]
Posterior probability of PDD > PD-MCI	37%	45%
Global cognitive ability (mm^2/sec/point)	1.2 [-0.7, 3.1]	-0.1 [-2, 1.7]
Posterior probability of a negative association	11%	56%
UPDRS-III (mm^2/sec/point)	0.4 [-0.6, 1.4]	-0.2 [-1.2, 0.8]
Posterior probability of a positive association	79%	35%

For controls, the reported value [95% uncertainty interval] represents the estimate for RD. The values are multiplied by 10^5. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: mm^2 per sec per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

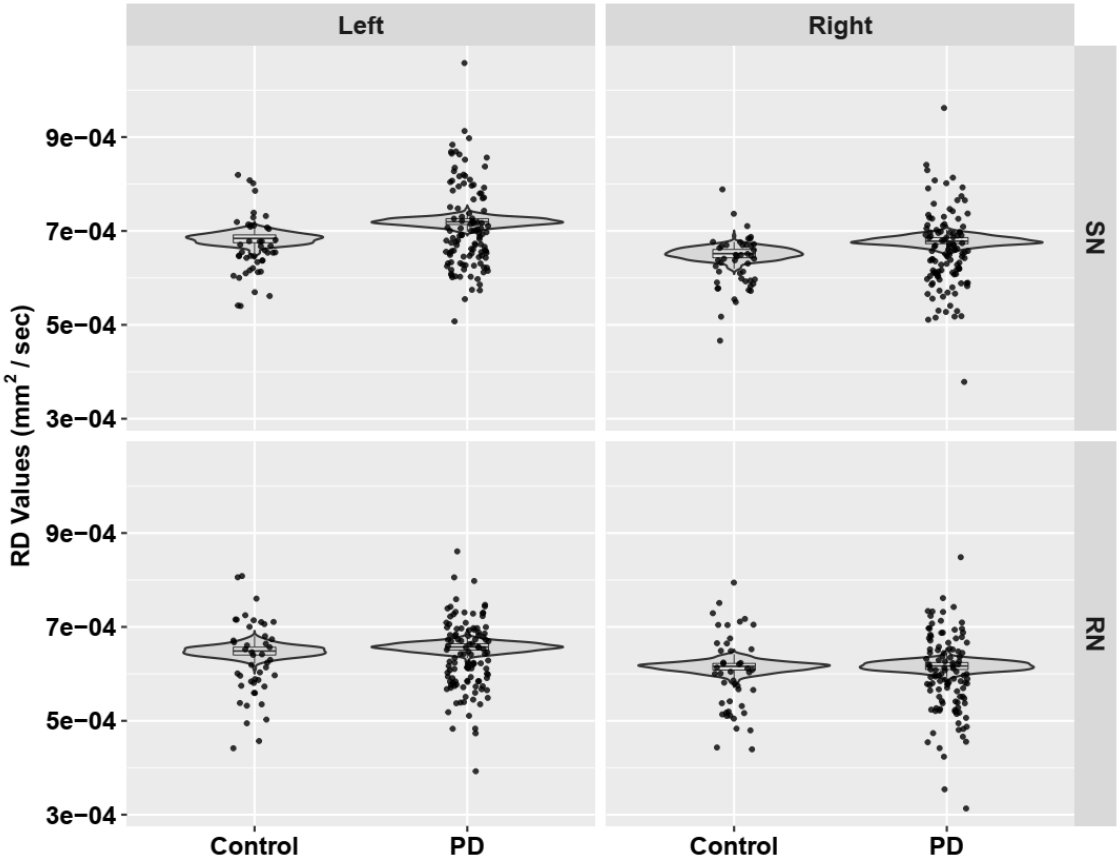


Figure 6.19: RD values of left and right sides of SN and RN between healthy control and PD. Dots are raw RD values. Violin plots in gray and boxplots represent the distribution of RD values adjusted for covariates from 200 draws from the posterior probability distribution.

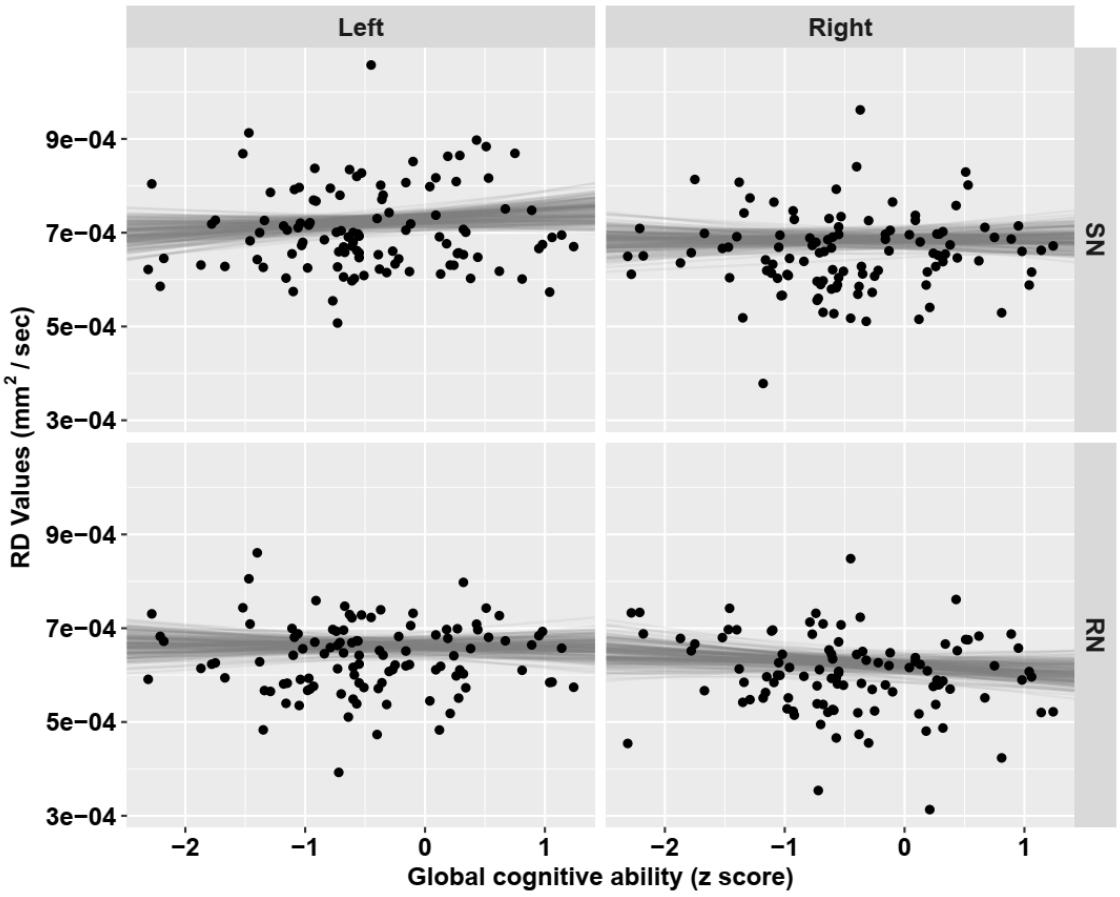


Figure 6.20: Correlations between RD values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

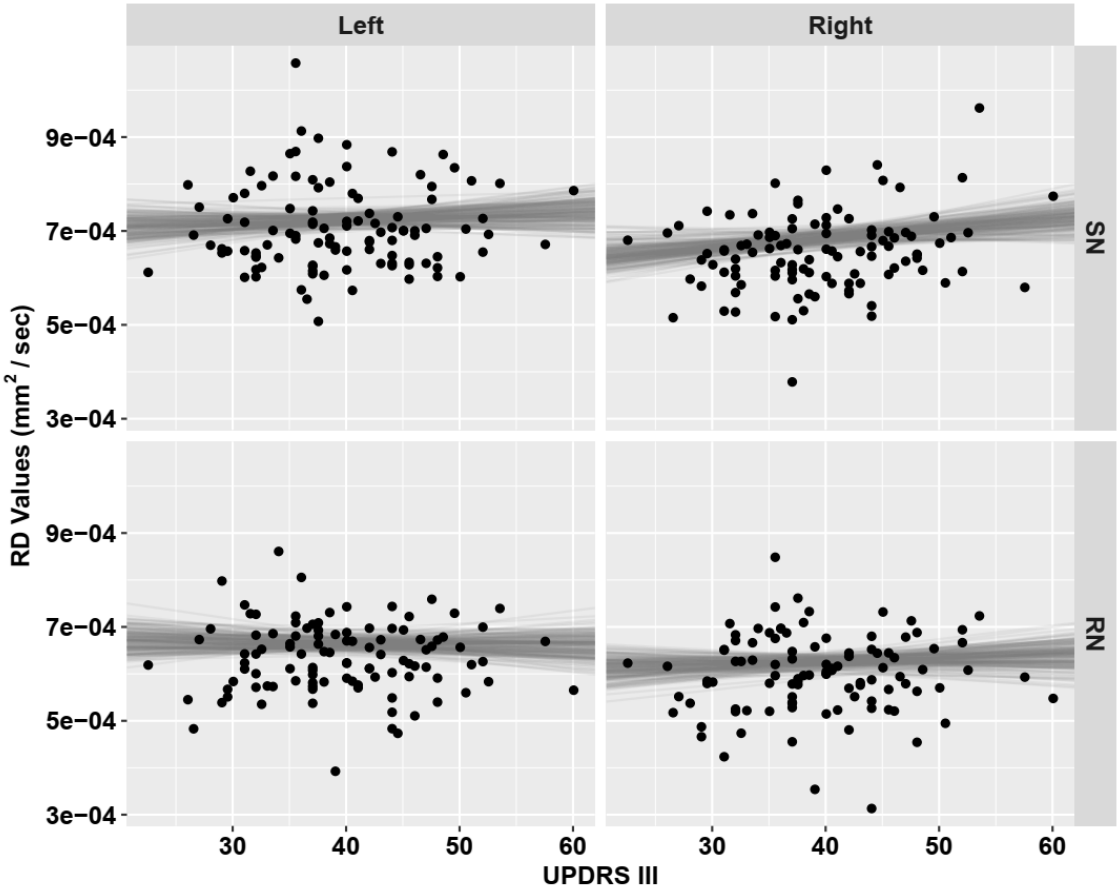


Figure 6.21: Correlations between RD values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

6.4.8 CBF

Table 6.9 and Figure 6.22 report the results from comparisons between healthy control and PD groups in SN and RN. We observed no evidence of decreased CBF values in the SN in the PD group relative to controls (probability P = 63%), and within the RN (P = 71%). Furthermore, we found moderate evidence for an association with age in the SN (probability P = 88% of a negative association), and in the RN (P = 98%). We also did not find evidence of a difference between anatomic side (probability P = 31% probability of right side having greater CBF values).

Table 6.9 also reports the results from comparisons between PD-N, PD-MCI, PDD and healthy control. We observed no evidence of decreased CBF values in the SN in PD-N (P = 38%) and PD-MCI (P = 58%) relative to controls, and within RN (P = 65% (PD-N), P = 47% (PD-MCI)). Furthermore, we found no evidence of decreased CBF values in the SN in PD-MCI (P = 70%) relative to PD-N and in the RN (P = 31%). However, we found weak evidence of decreased CBF values in the SN in PDD relative to controls (P = 78%), PD-N (P = 84%), and PD-MCI (P = 75%) and moderate evidence within RN (P = 94%, P = 88%, and P = 95%).

Table 6.9, Figure 6.23 and Figure 6.24 report the relationship between CBF, global cognitive ability, and motor impairment. We found no evidence of a positive relationship between CBF and global cognitive z score in SN (57% probability of a positive association between CBF and global cognitive ability) and within RN (74%). However, we observed weak evidence for a negative association with motor impairment in SN (75%), and in the RN (79%).

Table 6.9: CBF differences between Parkinson's disease (PD-N, PD-MCI, PDD) and control participants, associations between CBF, global cognitive ability, and motor impairment

	CBF	
Region	SN	RN
Control group (ml/min/100g)	388 [349, 426]	394 [354, 434]
PD vs controls (ml/min/100g)	-6.5 [-44, 31]	-11 [-49, 27]
Posterior probability of PD < controls	63%	71%
PD-N vs controls (ml/min/100g)	7.5 [-37, 54]	-9.5 [-57, 37]
Posterior probability of PD-N < controls	38%	65%
PD-MCI vs controls (ml/min/100g)	-4.2 [-45, 37]	1.5 [-41, 44]
Posterior probability of PD-MCI < controls	58%	47%
PDD vs controls (ml/min/100g)	-25 [-88, 35]	-49 [-113, 14]
Posterior probability of PDD < controls	78%	94%
PD-MCI vs PD-N (ml/min/100g)	-12 [-54, 31]	11 [-34, 55]
Posterior probability of PD-MCI < PD-N	70%	31%
PDD vs PD-N (ml/min/100g)	-32 [-97, 31]	-40 [-107, 26]
Posterior probability of PDD < PD-N	84%	88%
PDD vs PD-MCI (ml/min/100g)	-21 [-81, 40]	-51 [-113, 11]
Posterior probability of PDD < PD-MCI	75%	95%
Global cognitive ability (ml/min/100g/point)	2.5 [-23, 28]	9 [-17, 36]
Posterior probability of a positive association	57%	74%
UPDRS-III (ml/min/100g/point)	-5.2 [-20, 9.5]	-6.4 [-22, 9.3]
Posterior probability of a negative association	75%	79%

For controls, the reported value [95% uncertainty interval] represents the estimate for CBF. The values are multiplied by 10. For the comparisons, reported values represent estimated pairwise difference between specific groups and for the regressions with global cognitive ability and UPDRS-III, the reported value represents the slope of the association: ml per min per 100g per one point in global cognitive z score/UPDRS-III, accounting for age, sex, and anatomic side.

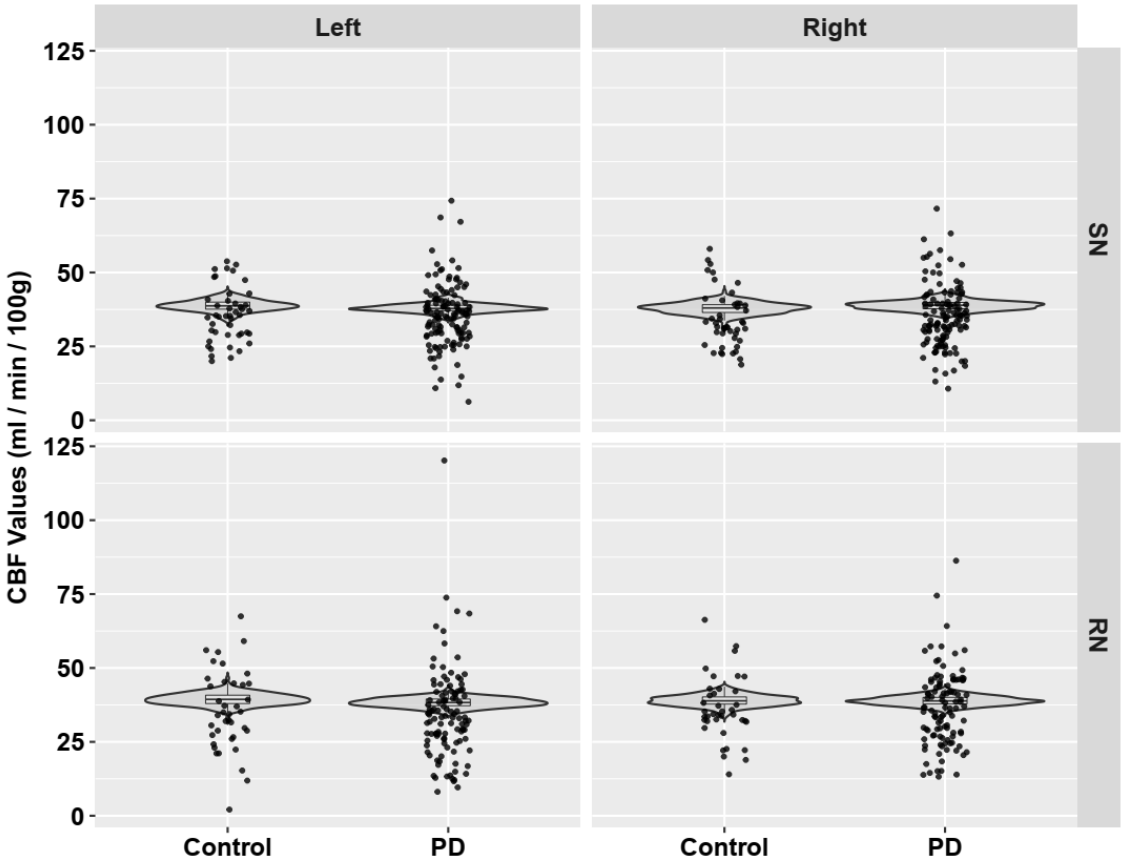


Figure 6.22: CBF values of left and right sides of SN and RN between healthy control and PD. Dots are raw CBF values. Violin plots in gray and boxplots represent the distribution of CBF values adjusted for covariates from 200 draws from the posterior probability distribution.

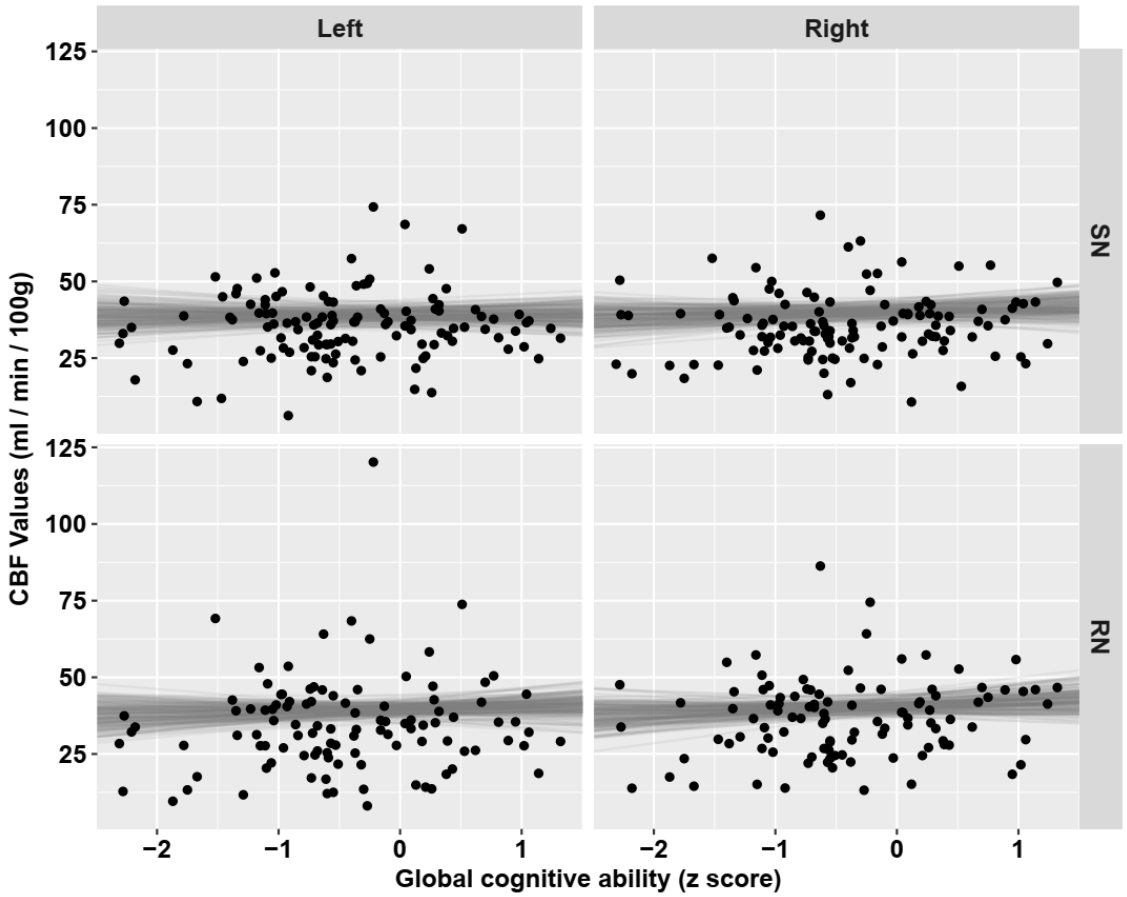


Figure 6.23: Correlations between CBF values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive z score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

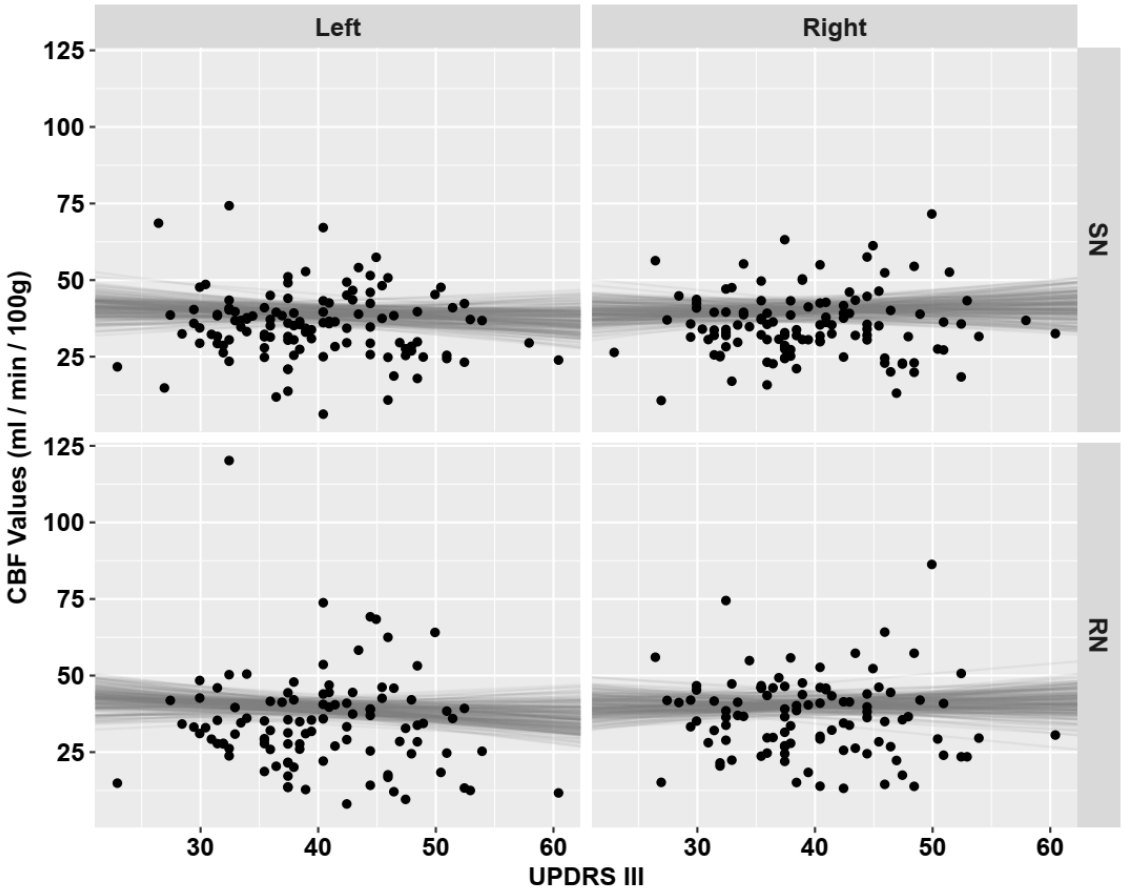


Figure 6.24: Correlations between CBF values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray.

6.4.9 Classification of PD and Healthy Controls

The receiver operating characteristics curve and the precision-recall curve for the linear SVM is shown in Figure 6.25. The AUCROC and AUCPR for linear SVM were 0.76 and 0.89, respectively, which were satisfactory. The AUCROC and AUCPR obtained by selecting the five most important features (QSM Right SN, R2* Left SN, R2* Left RN, R2* Right RN and R2* Right SN), obtained using feature ranking for linear SVM, were 0.72 and 0.7 respectively. The number of ranked featured used were selected at random. The results of the classification of PD and healthy controls using linear SVM and other classifiers including non-linear SVM, logistic regression, and random forest is shown in Table 6.10.

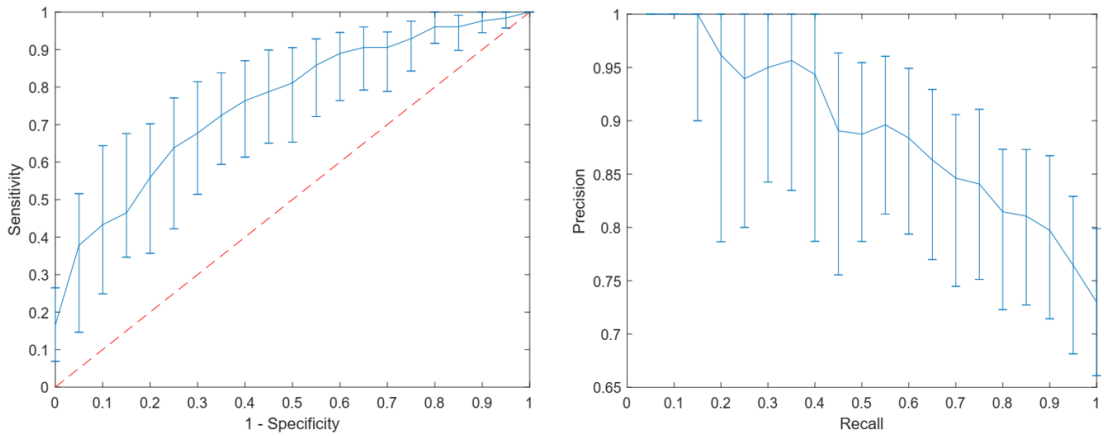


Figure 6.25: ROC curve and PR curve for linear SVM

Table 6.10: Classifier performance for predicting healthy controls and PD

Classifier	Sensitivity	Specificity	Precision	AUCROC	AUCPR
Linear SVM	0.64	0.77	0.88	0.76 [0.68, 0.83]	0.89 [0.83, 0.93]
Non-Linear SVM	0.62	0.66	0.83	0.66 [0.58, 0.75]	0.84 [0.76, 0.90]
Logistic Regression	0.68	0.70	0.86	0.74 [0.66, 0.81]	0.88 [0.82, 0.93]
Random Forest	0.93	0.30	0.78	0.68 [0.57, 0.76]	0.81 [0.71, 0.87]

AUCROC and AUCPR: The reported value [95% confidence interval]. SVM = Support vector machine, AUCROC = Area under the receiver operating characteristics curve, AUCPR = Area under the precision-recall curve.

6.5 Discussion

In this study, we assessed a number of different MRI metrics extracted from the SN and RN across a large cohort of healthy controls and PD participants. While some metrics showed strong evidence of differences in PD, there were metrics that did not exhibit this evidence. Key findings showed that quantitative susceptibility mapping is more sensitive in detecting tissue changes than volumetric measurements. QSM and R2* both allow *in vivo* investigation of iron accumulation in brain tissue. Our findings of strong evidence of increased QSM and R2* values in the SN in the PD group and moderate evidence in the RN, are consistent with the expected iron accumulation in PD patients (An et al., 2018; Barbosa et al., 2015; Chen et al., 2019; Du et al., 2016; Murakami et al., 2015; J. Y. Wang et al., 2016). As the SN is one of the earliest and most important structures to be affected in PD, it may be more vulnerable to the deposition of iron than the other cerebral regions. We also found strong evidence of increased QSM and R2* values in PD-N, PD-MCI and PDD in the SN compared to healthy controls, while we found moderate evidence of increased QSM in RN. A recent study by Uchida et al. (Uchida et al., 2019) also reported higher QSM in PD-N in SN compared to healthy controls. However, we did not find evidence of increased QSM and R2* in PD-MCI and PDD in SN and RN compared to PD-N but observed moderate evidence of increased QSM and R2* in PDD in SN compared to PD-MCI group. Furthermore, we also did not

find strong association between QSM and R2* values within SN and RN and either global cognitive ability or motor impairment, although we saw moderate evidence (a trend) of association between QSM in SN and cognition and QSM in RN and motor impairment. Overall, results from QSM and R2* suggest that iron deposition may not track cognitive or motor progression in PD and may therefore be more appropriate as a marker of the presence of PD (as opposed to a marker of disease progression). The results obtained in this study provide further support for the advantage of QSM over R2* in measuring iron content in brain tissues, possibly because it is capable of capturing purer susceptibility as it reflects the quantitative susceptibility of local tissues, while R2* captures the susceptibility of both local and surrounding tissues.

The QSM results presented in this chapter were analyzed across groups (both PD/control and across cognitive subgroups), global cognitive ability, and motor impairments in PD on a large cohort of the participants to possibly capture the full range of QSM and midbrain nuclei (SN and RN) variability during disease progression compared to the results presented in Chapter 5.

Contrary to initial hypotheses, we observed no evidence of reduction of volume in PD in the SN and RN—in fact, it appeared that PD exhibited volume increases relative to controls. While at first contradictory, these results are consistent with the method used to identify and define the SN. Increased iron deposition in the SN can lead to 'blooming' or extremely increased signal intensity on the QSM; this may extend beyond or bleed into surrounding tissue on the QSM map. As the SN was defined on the QSM images, the highly iron-laden SN was identified as being larger in volume, most likely due to increased iron accumulation. Volume may not necessarily be higher in PD; it may be reflecting a bias caused by higher levels of iron accumulation. However, it should be noted that we did not account for intracranial volume for analysis; only age, sex, and anatomic side were accounted.

DTI parameters allow *in vivo* investigation of microstructural changes in brain tissue. We observed moderate evidence of decreased FA and moderate-strong evidence of increased diffusivity in SN, consistent with a number of previous studies (Knossalla et al., 2018; Loane et al., 2016; Péran et al., 2010; Safai et al., 2020; Wei et al., 2016). These results are consistent with the expected microstructural changes in PD patients; our simple interpretation of neurodegeneration suggests that the process disrupts diffusion barriers. Moreover, we found moderate-weak evidence of increased MD and

RD in RN in PD. However, the DTI changes reported in the literature have not been very consistent. Some studies reported increased FA (Lenfeldt, Larsson, Nyberg, Birgander, & Forsgren, 2015) while others reported no changes in anisotropy in SN (Pelizzari et al., 2019; Safai et al., 2020) in PD. These inconsistencies may be due to varying methodologies, but they may be related to the limitations of the DTI model, i.e., free water (FW) contamination of the DTI measures (Ofori, Pasternak, Planetta, Burciu, et al., 2015; Planetta et al., 2016). Therefore, studies (Guttuso et al., 2018; Ofori, Pasternak, Planetta, Li, et al., 2015) have suggested FW assessments may provide more consistent results.

We also showed moderate evidence of decreased FA in PD-N, and weak evidence in PD-MCI in SN relative to controls. We found moderate evidence of increased AD in PD-N, PD-MCI, and PDD in SN compared to controls and weak evidence in PD-MCI in RN. Furthermore, strong-moderate evidence of increased MD, and strong-moderate-weak evidence of increased RD was observed in PD-N, PD-MCI and PDD in SN while moderate-weak evidence of increased MD was found in RN in PD-N and PD-MCI group. PD-N also exhibited moderate evidence of increased RD in RN compared to healthy controls. However, we did not observe evidence of decreased anisotropy and increased diffusivity in PD-MCI relative to PD-N, PDD-relative to PD-MCI and PDD relative to PD-N.

AD values within SN and RN exhibited weak evidence of an association with global cognitive ability while FA, MD and RD showed no evidence. We also demonstrated weak association in AD and RD values in the SN with motor impairment, while MD values exhibited moderate association. However, we did not find association of DTI parameters in RN with motor impairment. It is worth mentioning that two recent studies (atlas-based SN identification) (Pelizzari et al., 2019; Safai et al., 2020) did not perform correlations with a global cognitive score (the Montreal Cognitive Assessment) and/or the UPDRS-III. The study conducted by Apoorva et al., did not report significant association between DTI measures and UPDRS-III, which is in contrast to our study. UPDRS-III scores reported in their study were evaluated in the OFF-state while our cohort was assessed without interruption to participants' drug regime (i.e., most were assessed and scanned ON-state, but a few were 'drug naive', having not started yet anti-parkinsonian medications. Moreover, their study did not differentiate the PD patients into PD-N, PD-MCI and PDD, and hence the stage of the severity of the disease in the participants is not known which could overall drive the difference.

We observed contradictory results in SN and RN FA values with global cognitive ability. The hypothesis showed no evidence of positive association with cognition—in fact, it appeared that the PD exhibited negative association with global cognitive ability. Investigating the SN with free water diffusion or more advanced diffusion protocols may provide more information on this slightly surprising finding.

We also included relative motion in all the DTI comparisons in this study in an attempt to account for the effect of head motion on DTI metrics (Ling et al., 2012). In this cohort, inclusion of head motion parameters had minimal influence on results.

The DTI results reported in this study provide evidence of structural damage in SN and RN in PD patients. These results suggest that the tissue within the SN is also losing integrity, which may reflect the known dopaminergic cell death and loss of striatal projections.

We also assessed perfusion using ASL. We did not find evidence of a difference in CBF between PD and controls in SN and RN, which is in line with the study conducted by Pelizzari et al. (Pelizzari et al., 2019) and Al-Bachari et al. (Al-Bachari et al., 2014). However, we demonstrated weak evidence of decreased CBF in PDD compared to controls, PD-N and PD-MCI in SN, and moderate evidence within RN, similar to the CBF alterations reported in a previous study (Melzer et al., 2011). Hypoperfusion may indicate reduced metabolism, an underlying structural loss of tissue, or may reflect loss of functional connectivity, among other potential causes. CBF values within SN and RN also exhibited association with motor impairment. However, more studies using consistent ASL acquisitions in well-characterized groups of PD participants are needed to confirm the widely heterogeneous results present in the literature.

The results of the ROC and PR showed that the combination of multiple modalities were successful in classifying PD and healthy controls. The proposed method of utilizing various MR parameters (QSM, R2*, FA, AD, MD, RD, CBF and volume) performed better in discriminating PD and controls than using just two MR parameters (the five most important features comprised of QSM and R2*) obtained from feature ranking using ReliefF further confirms that MR parameters provide different but complementary information. Therefore, multiple MR metrics are able to provide varied information about the ongoing pathological processes in PD, and in turn is more successful in predicting healthy controls and PD than using limited or a single MR metric.

We also compared the performance of linear SVM with non-linear SVM, logistic regression, and random forest as these classification methods are widely used in the literature (Barbosa et al., 2015; Du et al., 2016; Safai et al., 2020) and significant results of distinction between the groups were shown. The different classifiers used for comparison also showed competitive results in discriminating between the two groups, however, linear SVM performed best among them.

Our study has several advantages. First, the study was conducted in a large cohort of extensively phenotyped healthy controls and PD participants and is one of the very first studies performed with such a large number of participants utilizing automated segmentation (discussed in the earlier chapter). Most of the state-of-the-art method used either manual segmentations or the atlas based segmentation. Moreover, the segmentation method was specifically designed and validated to identify SN and RN. Therefore, the segmentations produced by the proposed method should be more accurate than most other studies. Second, this was one of, if not the first study to use varied imaging modalities including QSM, R2*, DTI, ASL, and volume capable of capturing a wide range of several ongoing processes in both SN and RN during PD progression, and correlate with both global cognitive ability and motor impairment. Third, the broader distribution of duration of diagnosis (1 to 25 years) of the PD participants included in our study enabled us to capture the association between QSM, R2*, DTI, ASL, volume changes, and disease severity.

Our study has some limitations. First, the slices of the QSM, which were used to segment SN and RN were 2 mm thick. In the future, perhaps thinner (1 mm) slices would yield even more accurate segmentations. Second, the diffusion is robust but 'old'. Newer diffusion acquisitions allow fitting of more advanced models. Perhaps these would yield even more information about the state of the SN/RN. The algorithm could be extended using multiple MRI to improve segmentation of SN and RN. Additionally, longitudinal studies could confirm the association of different MRI metrics with disease progression.

6.6 Chapter Summary

This chapter presented the experimental analysis of midbrain nuclei iron content, microstructural alterations, perfusion assessment and volume atrophy studied using QSM, R2*, FA, AD, MD, RD, CBF and volume respectively. The evaluation was performed with the help of large database provided by the New Zealand Brain Research Institute consisting a total of 174 participants including both healthy controls and PD (PD-N, PD-MCI, and PDD) patients.

The study performed in this chapter showed strong-moderate evidence of increased QSM, R2*, AD, MD, and RD values and decreased FA values in the SN in PD relative to controls; QSM, R2* and MD values in RN showed moderate evidence of increment. We also exhibited consistent increment in QSM, R2*, AD, and MD values in SN in PD-N, PD-MCI and PDD compared to controls while decrement in FA was observed solely in PD-N. On the other hand, QSM, R2*, MD and RD values in RN were increased in PD-N compared to healthy controls. We also demonstrated moderate evidence of an association between MD values in SN with motor impairment. However, AD, RD, and CBF values within SN only showed weak evidence of an association. Moreover, just AD values within SN and RN exhibited weak evidence of an association with global cognitive ability. Linear SVM showed that the combination of different modalities were successful in discriminating PD and controls. Therefore, work presented in this chapter suggests that multiple imaging modalities measuring complementary tissue characteristics may be useful for monitoring several ongoing processes in midbrain nuclei in Parkinson's disease and also could be helpful for distinction between PD and controls.

7 CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

This thesis proposed robust algorithms for the segmentation of substantia nigra and red nucleus from MRI images. The segmentation of SN based on weighted thresholding approach employed *k*-means algorithm on different samples of the image to generate accurate segmentations on T2*-weighted images. Similarly, the methodology based on the combination of level set method and dual-tree complex wavelet transform was also designed to achieve better, effective, and robust segmentation of SN using T2*-weighted images.

This thesis also proposed an improved automated segmentation method based on substantia nigra seeding, contrast enhancement, level set method, DT-CWT and watershed transform for accurate, effective, and efficient segmentation of substantia nigra and red nucleus from QSM images; this framework was an improvement to the method based on level set and wavelet transform.

This thesis validated all the proposed algorithms through subjective evaluation and quantitative assessment performed based on manual segmentations produced by expert radiologists'. Moreover, segmentations of substantia nigra and red nucleus from QSM images facilitated *in vivo* assessment of the midbrain nuclei iron content, which is significant since increased iron accumulation in the smaller nuclei may play a role in the pathogenesis of Parkinson's disease. Specifically, this thesis showed increased QSM

values in SN in PD relative to controls, with moderate evidence for association between QSM, global cognitive ability, and motor impairment in the SN in PD, and between QSM and motor impairment in RN in PD. The results were comparable between the proposed segmentation and radiologists-drawn manual segmentation.

After validating the segmentation algorithm, the proposed framework was implemented on a large cohort of healthy controls and those with PD to perform a multi-modal examination including QSM, R2*, fractional anisotropy, mean diffusivity, axial diffusivity, radial diffusivity, cerebral blood flow and volume of substantia nigra and red nucleus across groups (both PD/control and across cognitive subgroups) and clinical scores, and interesting observations of associations were made including some strong evidences reported between groups. Furthermore, linear SVM classifier used with multi-modal MRI metrics from SN and RN (both left and right) as a feature set resulted in good accuracy in classifying PD from controls. Therefore, multi-modal MRI may provide a window into the brain changes associated with PD progression, as well as classify PD and healthy controls, and hence should be an active area of research in the future.

7.2 Future Work

This thesis provides ample possibilities for future research in the area of segmentation of midbrain nuclei, not just to explore and characterize Parkinson's disease, but potentially other conditions, as well. The segmentation approach proposed in this thesis could be extended to achieve finer-grain segmentation, for example segmentation of the pars compacta and pars reticulata individually, in order to study their individual role in PD. Moreover, SN segmentation into dorsal and ventral regions could also be performed to investigate their contribution to the pathogenesis of PD. The proposed segmentation algorithm is primarily focused on the substantia nigra and red nucleus; however, extension of the current method could be made for different brain regions, including other basal ganglia structures. Similarly, the proposed segmentation method could be applied to other disease groups, such as those with Alzheimer's disease.

Additionally, the segmentation results obtained from the proposed method could very well be improved with the application of cutting-edge techniques such as deep learning and by incorporating multiple imaging modalities into the segmentation

routines. Furthermore, newer imaging techniques offering stronger contrast of midbrain structures (e.g., neuromelanin sensitive MRI) may further improve segmentation.

The comprehensive multi-modal MRI examination of the substantia nigra and red nucleus, including QSM, R2*, DTI parameters, ASL, and volume measuring iron deposition, microstructural alterations, perfusion changes, and volume, respectively, presents an opportunity to increase our understanding of these two key structures related to Parkinson's disease. Results, meanwhile, raise the question of how these MRI values are affected in early PD (or even pre-clinical) as well as longitudinally over time within the same individuals. Moreover, the predictive analysis in classifying Parkinson's disease and healthy controls utilizing multi-modal MRI employing machine learning algorithms should also be further explored, and with the application of advanced feature ranking techniques as well as utilizing other basal ganglia structures for improved prediction.

8 REFERENCES

Acosta-Cabronero, J., Cardenas-Blanco, A., Betts, M. J., Butryn, M., Valdes-Herrera, J. P., Galazky, I., & Nestor, P. J. (2017, Jan). The whole-brain pattern of magnetic susceptibility perturbations in Parkinson's disease. *Brain, 140*(1), 118-131. https://doi.org/10.1093/brain/aww278

Aguirre, G. K., Detre, J. A., & Wang, J. (2005). Perfusion fMRI for functional neuroimaging. *Int Rev Neurobiol, 66*, 213-236. https://doi.org/10.1016/S0074-7742(05)66007-2

Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017, 2017/08/01). Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of Digital Imaging, 30*(4), 449-459. https://doi.org/10.1007/s10278-017-9983-4

Al-Bachari, S., Parkes, L. M., Vidyasagar, R., Hanby, M. F., Tharaken, V., Leroi, I., & Emsley, H. C. (2014). Arterial spin labelling reveals prolonged arterial arrival time in idiopathic Parkinson's disease. *Neuroimage Clin, 6*, 1-8. https://doi.org/10.1016/j.nicl.2014.07.014

Alkemade, A., de Hollander, G., Keuken, M. C., Schäfer, A., Ott, D. V. M., Schwarz, J., . . . Forstmann, B. U. (2017). Comparison of T2*-weighted and QSM contrasts in Parkinson's disease to visualize the STN with MRI. *Plos One, 12*(4), e0176130. https://doi.org/10.1371/journal.pone.0176130

Alsop, D. C., Detre, J. A., Golay, X., Gunther, M., Hendrikse, J., Hernandez-Garcia, L., . . . Zaharchuk, G. (2015, Jan). Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia. *Magnetic Resonance in Medicine, 73*(1), 102-116. https://doi.org/10.1002/mrm.25197

An, H., Zeng, X., Niu, T., Li, G., Yang, J., Zheng, L., . . . Li, J. (2018, 2018/03/15/). Quantifying iron deposition within the substantia nigra of Parkinson's disease by quantitative susceptibility mapping. *Journal of the Neurological Sciences, 386*, 46-52. https://doi.org/https://doi.org/10.1016/j.jns.2018.01.008

Aquino, D., Contarino, V., Albanese, A., Minati, L., Farina, L., Grisoli, M., . . . Chiapparini, L. (2014, 2014/05/01). Substantia nigra in Parkinson's disease: a multimodal MRI comparison between early and advanced stages of the disease. *Neurological Sciences, 35*(5), 753-758. https://doi.org/10.1007/s10072-013-1595-2

Ashburner, J., & Friston, K. J. (2005, Jul 1). Unified segmentation. *Neuroimage, 26*(3), 839-851. https://doi.org/10.1016/j.neuroimage.2005.02.018

Barbosa, J. H., Santos, A. C., Tumas, V., Liu, M., Zheng, W., Haacke, E. M., & Salmon, C. E. (2015, Jun). Quantifying brain iron deposition in patients with Parkinson's disease using quantitative susceptibility mapping, R2 and R2. *Magn Reson Imaging, 33*(5), 559-565. https://doi.org/10.1016/j.mri.2015.02.021

Basser, P. J. (1995, Nov-Dec). Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR Biomed, 8*(7-8), 333-344. https://doi.org/10.1002/nbm.1940080707

Basukala, D., Mukundan, R., Lim, A., Hurrell, M. A., Keenan, R. J., Dalrymple-Alford, J. C., . . . Melzer, T. R. (2021, 2021/05/01/). Automated segmentation of substantia nigra and red nucleus using quantitative susceptibility mapping images: Application to Parkinson's disease. *Computers & Electrical Engineering, 91*, 107091. https://doi.org/https://doi.org/10.1016/j.compeleceng.2021.107091

Basukala, D., Mukundan, R., Melzer, T., & Keenan, R. (2018a, 12-14 Dec. 2018). *Segmentation of Substantia Nigra for the Automated Characterization of Parkinson's Disease.* Paper presented at the 2018 IEEE International Conference on Image Processing, Applications and Systems (IPAS).

Basukala, D., Mukundan, R., Melzer, T., & Keenan, R. (2018b, 19-21 Nov. 2018). *Segmentation of Substantia Nigra Using Weighted Thresholding Method.* Paper presented at the 2018 International Conference on Image and Vision Computing New Zealand (IVCNZ).

Basukala, D., Mukundan, R., Melzer, T., & Lim, A. (2019, 5-7 Dec. 2019). *Automated Segmentation of Substantia Nigra and Red Nucleus in Quantitative Susceptibility Mapping Images.* Paper presented at the 2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT).

Bergsland, N., Zivadinov, R., Schweser, F., Hagemeier, J., Lichter, D., & Guttuso, T., Jr. (2019, Jul). Ventral posterior substantia nigra iron increases over 3 years in Parkinson's disease. *Mov Disord, 34*(7), 1006-1013. https://doi.org/10.1002/mds.27730

Bilgic, B., Pfefferbaum, A., Rohlfing, T., Sullivan, E. V., & Adalsteinsson, E. (2012, 2012/02/01/). MRI estimates of brain iron concentration in normal aging using quantitative susceptibility mapping. *Neuroimage, 59*(3), 2625-2635. https://doi.org/https://doi.org/10.1016/j.neuroimage.2011.08.077

Bjorklund, G., Hofer, T., Nurchi, V. M., & Aaseth, J. (2019, Oct). Iron and other metals in the pathogenesis of Parkinson's disease: Toxic effects and possible detoxification. *J Inorg Biochem, 199*, 110717. https://doi.org/10.1016/j.jinorgbio.2019.110717

Borgefors, G. (1986, Jun). Distance Transformations in Digital Images. *Computer Vision Graphics and Image Processing, 34*(3), 344-371. https://doi.org/Doi 10.1016/S0734-189x(86)80047-0

Brain, W. R. B., & Walton, J. N. (1977). *Brain's Diseases of the nervous system* (8th ed.). Oxford Eng. ; New York: Oxford University Press.

Bürkner, P.-C. (2017, 08/01). brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software, 80*. https://doi.org/10.18637/jss.v080.i01

Celik, T., & Tjahjadi, T. (2010, Jul). Unsupervised colour image segmentation using dual-tree complex wavelet transform. *Computer Vision and Image Understanding, 114*(7), 813-826. https://doi.org/10.1016/j.cviu.2010.03.002

Chan, T. F., & Vese, L. A. (2001, Feb). Active contours without edges. *Ieee Transactions on Image Processing, 10*(2), 266-277. https://doi.org/Doi 10.1109/83.902291

Chavhan, G. B., Babyn, P. S., Thomas, B., Shroff, M. M., & Haacke, E. M. (2009, Sep-Oct). Principles, techniques, and applications of T2*-based MR imaging and its special applications. *Radiographics, 29*(5), 1433-1449. https://doi.org/10.1148/rg.295095034

Chen, Q., Chen, Y., Zhang, Y., Wang, F., Yu, H., Zhang, C., . . . Luo, W. (2019, 2019/05/22). Iron deposition in Parkinson's disease by quantitative susceptibility mapping. *BMC Neuroscience, 20*(1), 23. https://doi.org/10.1186/s12868-019-0505-9

Cheng, J. R., & Rajapakse, J. C. (2009, Mar). Segmentation of Clustered Nuclei With Shape Markers and Marking Function. *Ieee Transactions on Biomedical Engineering, 56*(3), 741-748. https://doi.org/10.1109/Tbme.2008.2008635

Churchill, N. W., Yourganov, G., & Strother, S. C. (2014). Comparing within-subject classification and regularization methods in fMRI for large and small sample sizes. *Hum Brain Mapp, 35*(9), 4499-4517. https://doi.org/10.1002/hbm.22490

Dai, W., Garcia, D., de Bazelaire, C., & Alsop, D. C. (2008). Continuous flow-driven inversion for arterial spin labeling using pulsed radio frequency and gradient fields. *Magnetic Resonance in Medicine, 60*(6), 1488-1497. https://doi.org/10.1002/mrm.21790

de Rochefort, L., Liu, T., Kressler, B., Liu, J., Spincemaille, P., Lebon, V., . . . Wang, Y. (2010, Jan). Quantitative susceptibility map reconstruction from MR phase data using bayesian regularization: validation and application to brain imaging. *Magnetic Resonance in Medicine, 63*(1), 194-206. https://doi.org/10.1002/mrm.22187

Deistung, A., Schäfer, A., Schweser, F., Biedermann, U., Turner, R., & Reichenbach, J. R. (2013, 2013/01/15/). Toward in vivo histology: A comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R2∗-imaging at ultra-high magnetic field strength. *Neuroimage, 65*, 299-314. https://doi.org/https://doi.org/10.1016/j.neuroimage.2012.09.055

Detre, J. A., Leigh, J. S., Williams, D. S., & Koretsky, A. P. (1992, Jan). Perfusion imaging. *Magnetic Resonance in Medicine, 23*(1), 37-45. https://doi.org/10.1002/mrm.1910230106

Detre, J. A., Wang, J., Wang, Z., & Rao, H. (2009, Aug). Arterial spin-labeled perfusion MRI in basic and clinical neuroscience. *Curr Opin Neurol, 22*(4), 348-355. https://doi.org/10.1097/WCO.0b013e32832d9505

Dice, L. R. (1945). Measures of the Amount of Ecologic Association between Species. *Ecology, 26*(3), 297-302. https://doi.org/Doi 10.2307/1932409

Ding, Y., Acosta, R., Enguix, V., Suffren, S., Ortmann, J., Luck, D., . . . Lodygensky, G. A. (2020, 2020-March-26). Using Deep Convolutional Neural Networks for Neonatal Brain Image Segmentation. *Frontiers in Neuroscience, 14*(207). https://doi.org/10.3389/fnins.2020.00207

Dong, J., Liu, T., Chen, F., Zhou, D., Dimov, A., Raj, A., . . . Wang, Y. (2015). Simultaneous phase unwrapping and removal of chemical shift (SPURS) using graph cuts: application in quantitative susceptibility mapping. *Ieee Transactions on Medical Imaging, 34*(2), 531-540. https://doi.org/10.1109/TMI.2014.2361764

Dorsey, E. R., Constantinescu, R., Thompson, J. P., Biglan, K. M., Holloway, R. G., Kieburtz, K., . . . Tanner, C. M. (2007, Jan 30). Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology, 68*(5), 384-386. https://doi.org/DOI 10.1212/01.wnl.0000247740.47667.03

Drayer, B., Burger, P., Darwin, R., Riederer, S., Herfkens, R., & Johnson, G. A. (1986, Jul). MRI of brain iron. *AJR Am J Roentgenol, 147*(1), 103-110. https://doi.org/10.2214/ajr.147.1.103

Du, G., Lewis, M. M., Sica, C., He, L., Connor, J. R., Kong, L., . . . Huang, X. (2018, Sep). Distinct progression pattern of susceptibility MRI in the substantia nigra of Parkinson's patients. *Mov Disord, 33*(9), 1423-1431. https://doi.org/10.1002/mds.27318

Du, G., Lewis, M. M., Styner, M., Shaffer, M. L., Sen, S., Yang, Q. X., & Huang, X. (2011, Aug 1). Combined R2* and diffusion tensor imaging changes in the substantia nigra in Parkinson's disease. *Mov Disord, 26*(9), 1627-1632. https://doi.org/10.1002/mds.23643

Du, G., Lewis, M. M., Styner, M., Shaffer, M. L., Sen, S., Yang, Q. X., & Huang, X. (2011). Combined R2* and diffusion tensor imaging changes in the substantia nigra in Parkinson's disease. *Mov Disord, 26*(9), 1627-1632. https://doi.org/10.1002/mds.23643

Du, G., Liu, T., Lewis, M. M., Kong, L., Wang, Y., Connor, J., . . . Huang, X. (2016, Mar). Quantitative susceptibility mapping of the midbrain in Parkinson's disease. *Mov Disord, 31*(3), 317-324. https://doi.org/10.1002/mds.26417

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., . . . Dubois, B. (2007, Sep 15). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov Disord, 22*(12), 1689-1707; quiz 1837. https://doi.org/10.1002/mds.21507

Fearnley, J. M., & Lees, A. J. (1991, Oct). Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain, 114 (Pt 5)*, 2283-2301.

Forstmann, B. U., Keuken, M. C., Schafer, A., Bazin, P.-L., Alkemade, A., & Turner, R. (2014a, 12/09/online). Multi-modal ultra-high resolution structural 7-Tesla MRI data repository. *Scientific Data, 1*, 140050. https://doi.org/10.1038/sdata.2014.50

Forstmann, B. U., Keuken, M. C., Schafer, A., Bazin, P.-L., Alkemade, A., & Turner, R. (2014b). NITRC. Retrieved from https://www.nitrc.org/projects/atag_mri_scans/

Garzón, B., Sitnikov, R., Bäckman, L., & Kalpouzos, G. (2018, Apr 15). Automated segmentation of midbrain structures with high iron content. *Neuroimage, 170*, 199-209. https://doi.org/10.1016/j.neuroimage.2017.06.016

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., . . . Movement Disorder Society, U. R. T. F. (2008, Nov 15). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord, 23*(15), 2129-2170. https://doi.org/10.1002/mds.22340

Grau, V., Mewes, A. U. J., Alcaniz, M., Kikinis, R., & Warfield, S. K. (2004, Apr). Improved watershed transform for medical image segmentation using prior information. *Ieee Transactions on Medical Imaging, 23*(4), 447-458. https://doi.org/10.1109/Tmi.2004.824224

Grøvik, E., Yi, D., Iv, M., Tong, E., Nilsen, L. B., Latysheva, A., . . . Zaharchuk, G. (2021, 2021/02/22). Handling missing MRI sequences in deep learning segmentation of brain metastases: a multicenter study. *npj Digital Medicine, 4*(1), 33. https://doi.org/10.1038/s41746-021-00398-4

Guo, T., Song, Y., Li, J., Fan, M., Yan, X., He, A., . . . Yang, G. (2018, Mar 31). Seed point discontinuity-based segmentation method for the substantia nigra and the red nucleus in quantitative susceptibility maps. *J Magn Reson Imaging*. https://doi.org/10.1002/jmri.26023

Guttuso, T., Jr., Bergsland, N., Hagemeier, J., Lichter, D. G., Pasternak, O., & Zivadinov, R. (2018, Mar). Substantia Nigra Free Water Increases Longitudinally in Parkinson Disease. *AJNR Am J Neuroradiol, 39*(3), 479-484. https://doi.org/10.3174/ajnr.A5545

Haacke, E. M., Liu, S., Buch, S., Zheng, W., Wu, D., & Ye, Y. (2015, Jan). Quantitative susceptibility mapping: current status and future directions. *Magn Reson Imaging, 33*(1), 1-25. https://doi.org/10.1016/j.mri.2014.09.004

Haegelen, C., Coupe, P., Fonov, V., Guizard, N., Jannin, P., Morandi, X., & Collins, D. L. (2013, Jan). Automated segmentation of basal ganglia and deep brain structures in MRI of Parkinson's disease. *Int J Comput Assist Radiol Surg, 8*(1), 99-110. https://doi.org/10.1007/s11548-012-0675-8

Hallgren, B., & Sourander, P. (1958, Oct). The effect of age on the non-haemin iron in the human brain. *J Neurochem, 3*(1), 41-51.

Huang, Z.-K., & Chau, K.-W. (2008, 2008/11/15/). A new image thresholding method based on Gaussian mixture model. *Applied Mathematics and Computation, 205*(2), 899-907. https://doi.org/https://doi.org/10.1016/j.amc.2008.05.130

Hughes, A. J., Daniel, S. E., Kilford, L., & Lees, A. J. (1992, Mar). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry, 55*(3), 181-184. https://doi.org/10.1136/jnnp.55.3.181

Jain, A. K. (2010, 2010/06/01/). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*(8), 651-666. https://doi.org/https://doi.org/10.1016/j.patrec.2009.09.011

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012, Aug 15). Fsl. *Neuroimage, 62*(2), 782-790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Johansen-Berg, H., & Behrens, T. E. (2006, Aug). Just pretty pictures? What diffusion tractography can add in clinical neuroscience. *Curr Opin Neurol, 19*(4), 379-385. https://doi.org/10.1097/01.wco.0000236618.82086.01

Kalti, M. A. M. a. K. (2012). Image segmentation by adaptive distance based on EM algorithm. *CoRR*.

Khan, N. U., Arya, K. V., & Pattanaik, M. (2013, Jun). Histogram statistics based variance controlled adaptive threshold in anisotropic diffusion for low contrast image enhancement. *Signal Processing, 93*(6), 1684-1693. https://doi.org/10.1016/j.sigpro.2012.09.009

Kim, J., Lenglet, C., Duchin, Y., Sapiro, G., & Harel, N. (2014, Sep). Semiautomatic segmentation of brain subcortical structures from high-field MRI. *IEEE J Biomed Health Inform, 18*(5), 1678-1695. https://doi.org/10.1109/JBHI.2013.2292858

Kingsbury, N. (1998). *The Dual-Tree Complex Wavelet Transform: A New Technique For Shift Invariance And Directional Filters.* Paper presented at the 8th IEEE DSP Workshop, Utah.

Kingsbury, N. (2001, 2001/05/01/). Complex Wavelets for Shift Invariant Analysis and Filtering of Signals. *Applied and Computational Harmonic Analysis, 10*(3), 234-253. https://doi.org/https://doi.org/10.1006/acha.2000.0343

Kingsley, P. B. (2006). Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors. *Concepts in Magnetic Resonance Part A, 28A*(2), 101-122. https://doi.org/https://doi.org/10.1002/cmr.a.20048

Knossalla, F., Kohl, Z., Winkler, J., Schwab, S., Schenk, T., Engelhorn, T., . . . Gölitz, P. (2018, 2018/04/01/). High-resolution diffusion tensor-imaging indicates asymmetric microstructural disorganization within substantia nigra in early Parkinson's disease. *Journal of Clinical Neuroscience*, *50*, 199-202. https://doi.org/https://doi.org/10.1016/j.jocn.2018.01.023

Koo, T. K., & Li, M. Y. (2016, Jun). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Laszlo, M., & Mukherjee, S. (2006). A genetic algorithm using hyper-quadtrees for low-dimensional k-means clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(4), 533-543. https://doi.org/10.1109/TPAMI.2006.66

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001, Apr). Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging*, *13*(4), 534-546. https://doi.org/10.1002/jmri.1076

Lenfeldt, N., Larsson, A., Nyberg, L., Birgander, R., & Forsgren, L. (2015, Oct). Fractional anisotropy in the substantia nigra in Parkinson's disease: a complex picture. *Eur J Neurol*, *22*(10), 1408-1414. https://doi.org/10.1111/ene.12760

Li, C., Huang, R., Ding, Z., Gatenby, J. C., Metaxas, D. N., & Gore, J. C. (2011, Jul). A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Trans Image Process*, *20*(7), 2007-2016. https://doi.org/10.1109/TIP.2011.2146190

Li, C., Kao, C. Y., Gore, J. C., & Ding, Z. (2008, Oct). Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans Image Process*, *17*(10), 1940-1949. https://doi.org/10.1109/TIP.2008.2002304

Li, C. M., Kao, C. Y., Gore, J. C., & Ding, Z. H. (2007). Implicit active contours driven by local binary fitting energy. *2007 Ieee Conference on Computer Vision and Pattern Recognition, Vols 1-8*, 339-+.

Ling, J., Merideth, F., Caprihan, A., Pena, A., Teshiba, T., & Mayer, A. R. (2012, Jan). Head injury or head motion? Assessment and quantification of motion artifacts in diffusion tensor imaging studies. *Hum Brain Mapp*, *33*(1), 50-62. https://doi.org/10.1002/hbm.21192

Litvan, I., Goldman, J. G., Tröster, A. I., Schmand, B. A., Weintraub, D., Petersen, R. C., . . . Emre, M. (2012, Mar). Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Mov Disord*, *27*(3), 349-356. https://doi.org/10.1002/mds.24893

Liu, C., Li, W., Tong, K. A., Yeom, K. W., & Kuzminski, S. (2015, Jul). Susceptibility-weighted imaging and quantitative susceptibility mapping in the brain. *J Magn Reson Imaging*, *42*(1), 23-41. https://doi.org/10.1002/jmri.24768

Liu, D. J., & Yu, J. A. (2009). Otsu method and K-means. *His 2009: 2009 Ninth International Conference on Hybrid Intelligent Systems, Vol 1, Proceedings*, 344-349. https://doi.org/10.1109/His.2009.74

Liu, J., & Guo, L. (2015, 26-27 Aug. 2015). *A New Brain MRI Image Segmentation Strategy Based on K-means Clustering and SVM.* Paper presented at the 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics.

Liu, T., Eskreis-Winkler, S., Schweitzer, A. D., Chen, W., Kaplitt, M. G., Tsiouris, A. J., & Wang, Y. (2013, Oct). Improved subthalamic nucleus depiction with quantitative susceptibility mapping. *Radiology*, *269*(1), 216-223. https://doi.org/10.1148/radiol.13121991

Liu, T., Wisnieff, C., Lou, M., Chen, W. W., Spincemaille, P., & Wang, Y. (2013, Feb). Nonlinear formulation of the magnetic field to source relationship for robust quantitative susceptibility mapping. *Magnetic Resonance in Medicine*, *69*(2), 467-476. https://doi.org/10.1002/mrm.24272

Liu, T., Xu, W. Y., Spincemaille, P., Avestimehr, A. S., & Wang, Y. (2012, Mar). Accuracy of the Morphology Enabled Dipole Inversion (MEDI) Algorithm for Quantitative Susceptibility Mapping in MRI. *Ieee Transactions on Medical Imaging*, *31*(3), 816-824. https://doi.org/10.1109/Tmi.2011.2182523

Liu, Z., Spincemaille, P., Yao, Y. H., Zhang, Y., & Wang, Y. (2018, May). MEDI+0: Morphology enabled dipole inversion with automatic uniform cerebrospinal fluid zero reference for

quantitative susceptibility mapping. *Magnetic Resonance in Medicine, 79*(5), 2795-2803. https://doi.org/10.1002/mrm.26946

Loane, C., Politis, M., Kefalopoulou, Z., Valle-Guzman, N., Paul, G., Widner, H., . . . Piccini, P. (2016, Jul). Aberrant nigral diffusion in Parkinson's disease: A longitudinal diffusion tensor imaging study. *Mov Disord, 31*(7), 1020-1026. https://doi.org/10.1002/mds.26606

Lundervold, A. S., & Lundervold, A. (2019, 2019/05/01/). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik, 29*(2), 102-127. https://doi.org/https://doi.org/10.1016/j.zemedi.2018.11.002

Mallat, S. G. (1999). *A wavelet tour of signal processing* (2nd ed.). San Diego: Academic Press.

Martin, W. R., Wieler, M., & Gee, M. (2008, Apr 15). Midbrain iron content in early Parkinson disease: a potential biomarker of disease status. *Neurology, 70*(16 Pt 2), 1411-1417. https://doi.org/10.1212/01.wnl.0000286384.31050.b5

Mateos-Pérez, J. M., Dadar, M., Lacalle-Aurioles, M., Iturria-Medina, Y., Zeighami, Y., & Evans, A. C. (2018, 2018/01/01/). Structural neuroimaging as clinical predictor: A review of machine learning applications. *NeuroImage: Clinical, 20*, 506-522. https://doi.org/https://doi.org/10.1016/j.nicl.2018.08.019

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46. https://doi.org/10.1037/1082-989X.1.1.30

Melzer, T. R., Watts, R., MacAskill, M. R., Pearson, J. F., Rüeger, S., Pitcher, T. L., . . . Anderson, T. J. (2011). Arterial spin labelling reveals an abnormal cerebral perfusion pattern in Parkinson's disease. *Brain, 134*(3), 845-855. https://doi.org/10.1093/brain/awq377

Milletari, F., Ahmadi, S.-A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., . . . Navab, N. (2017, 2017/11/01/). Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Computer Vision and Image Understanding, 164*, 92-102. https://doi.org/https://doi.org/10.1016/j.cviu.2017.04.002

Mumford, D., & Shah, J. (1989, Jul). Optimal Approximations by Piecewise Smooth Functions and Associated Variational-Problems. *Communications on Pure and Applied Mathematics, 42*(5), 577-685. https://doi.org/DOI 10.1002/cpa.3160420503

Munoz, Y., Carrasco, C. M., Campos, J. D., Aguirre, P., & Nunez, M. T. (2016). Parkinson's Disease: The Mitochondria-Iron Link. *Parkinsons Dis, 2016*, 7049108. https://doi.org/10.1155/2016/7049108

Murakami, Y., Kakeda, S., Watanabe, K., Ueda, I., Ogasawara, A., Moriya, J., . . . Korogi, Y. (2015, Jun). Usefulness of quantitative susceptibility mapping for the diagnosis of Parkinson disease. *AJNR Am J Neuroradiol, 36*(6), 1102-1108. https://doi.org/10.3174/ajnr.A4260

Ng, H. P., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. (2006). Medical image segmentation using K-MEANS clustering and improved watershed algorithm. *7th Ieee Southwest Symposium on Image Analysis and Interpretation*, 61-+.

Noble, W. S. (2006, 2006/12//). What is a support vector machine? *Nature biotechnology, 24*(12), 1565-1567. https://doi.org/10.1038/nbt1206-1565

Oakley, A. E., Collingwood, J. F., Dobson, J., Love, G., Perrott, H. R., Edwardson, J. A., . . . Morris, C. M. (2007, May 22). Individual dopaminergic neurons show raised iron levels in Parkinson disease. *Neurology, 68*(21), 1820-1825. https://doi.org/10.1212/01.wnl.0000262033.01945.9a

Ofori, E., Pasternak, O., Planetta, P. J., Burciu, R., Snyder, A., Febo, M., . . . Vaillancourt, D. E. (2015, Feb). Increased free water in the substantia nigra of Parkinson's disease: a single-site and multi-site study. *Neurobiol Aging, 36*(2), 1097-1104. https://doi.org/10.1016/j.neurobiolaging.2014.10.029

Ofori, E., Pasternak, O., Planetta, P. J., Li, H., Burciu, R. G., Snyder, A. F., . . . Vaillancourt, D. E. (2015, Aug). Longitudinal changes in free-water within the substantia nigra of Parkinson's disease. *Brain, 138*(Pt 8), 2322-2331. https://doi.org/10.1093/brain/awv136

Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012, Apr). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev, 36*(4), 1140-1152. https://doi.org/10.1016/j.neubiorev.2012.01.004

Osher, S., & Sethian, J. A. (1988, 1988/11/01/). Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics, 79*(1), 12-49. https://doi.org/https://doi.org/10.1016/0021-9991(88)90002-2

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics, 9*(1), 62-66. https://doi.org/10.1109/TSMC.1979.4310076

Park, S.-C., Cha, J. H., Lee, S., Jang, W., Lee, C. S., & Lee, J. K. (2019, 2019-October-24). Deep Learning-Based Deep Brain Stimulation Targeting and Clinical Applications. *Frontiers in Neuroscience, 13*(1128). https://doi.org/10.3389/fnins.2019.01128

Pelizzari, L., Laganà, M. M., Di Tella, S., Rossetto, F., Bergsland, N., Nemni, R., . . . Baglio, F. (2019, 2019-June-04). Combined Assessment of Diffusion Parameters and Cerebral Blood Flow Within Basal Ganglia in Early Parkinson's Disease. *Frontiers in Aging Neuroscience, 11*(134). https://doi.org/10.3389/fnagi.2019.00134

Pelzer, E. A., Florin, E., & Schnitzler, A. (2019, 2019-August-06). Quantitative Susceptibility Mapping and Resting State Network Analyses in Parkinsonian Phenotypes—A Systematic Review of the Literature. *Frontiers in Neural Circuits, 13*(50). https://doi.org/10.3389/fncir.2019.00050

Péran, P., Cherubini, A., Assogna, F., Piras, F., Quattrocchi, C., Peppe, A., . . . Sabatini, U. (2010). Magnetic resonance imaging markers of Parkinson's disease nigrostriatal signature. *Brain, 133*(11), 3423-3433. https://doi.org/10.1093/brain/awq212

Planetta, P. J., Ofori, E., Pasternak, O., Burciu, R. G., Shukla, P., DeSimone, J. C., . . . Vaillancourt, D. E. (2016, Feb). Free-water imaging in Parkinson's disease and atypical parkinsonism. *Brain, 139*(Pt 2), 495-508. https://doi.org/10.1093/brain/awv361

Robnik-Šikonja, M., & Kononenko, I. (2003, 2003/10/01). Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning, 53*(1), 23-69. https://doi.org/10.1023/A:1025667309714

Roerdink, J. B. T. M., & Meijster, A. (2000). The Watershed Transform: Definitions, Algorithms and Parallelization Strategies. *Fundamenta Informaticae, 41*, 187-228. https://doi.org/10.3233/FI-2000-411207

Safai, A., Prasad, S., Chougule, T., Saini, J., Pal, P. K., & Ingalhalikar, M. (2020, Apr 1). Microstructural abnormalities of substantia nigra in Parkinson's disease: A neuromelanin sensitive MRI atlas based study. *Hum Brain Mapp, 41*(5), 1323-1333. https://doi.org/10.1002/hbm.24878

Schenck, J. F. (1996, Jun). The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Medical Physics, 23*(6), 815-850. https://doi.org/10.1118/1.597854

Schweser, F., Sommer, K., Deistung, A., & Reichenbach, J. R. (2012, Sep). Quantitative susceptibility mapping for investigating subtle susceptibility variations in the human brain. *Neuroimage, 62*(3), 2083-2100. https://doi.org/10.1016/j.neuroimage.2012.05.067

Sendra-Balcells, C., Salvador, R., Pedro, J. B., Biagi, M. C., Aubinet, C., Manor, B., . . . Ruffini, G. (2020). Convolutional neural network MRI segmentation for fast and robust optimization of transcranial electrical current stimulation of the human brain. *bioRxiv*, 2020.2001.2029.924985. https://doi.org/10.1101/2020.01.29.924985

Shan, S., Sandham, W., Granat, M., & Sterr, A. (2005). MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Transactions on Information Technology in Biomedicine, 9*(3), 459-467. https://doi.org/10.1109/TITB.2005.847500

Singh, A., Yadav, S., & Singh, N. (2016). Contrast Enhancement and Brightness Preservation using Global-Local Image Enhancement Techniques. *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (Pdgc)*, 291-294.

Sofic, E., Paulus, W., Jellinger, K., Riederer, P., & Youdim, M. B. (1991, Mar). Selective increase of iron in substantia nigra zona compacta of parkinsonian brains. *J Neurochem, 56*(3), 978-982.

Soille, P. (2003). *Morphological Image Analysis: Principles and Applications*: Springer-Verlag.

Thomas, G. E. C., Leyland, L. A., Schrag, A.-E., Lees, A. J., Acosta-Cabronero, J., & Weil, R. S. (2020). Brain iron deposition is linked with cognitive severity in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 91*(4), 418.

Uchida, Y., Kan, H., Sakurai, K., Arai, N., Kato, D., Kawashima, S., . . . Matsukawa, N. (2019). Voxel-based quantitative susceptibility mapping in Parkinson's disease with mild cognitive impairment. *Mov Disord, 34*(8), 1164-1173. https://doi.org/10.1002/mds.27717

Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018, Sep). Relief-based feature selection: Introduction and review. *J Biomed Inform, 85*, 189-203. https://doi.org/10.1016/j.jbi.2018.07.014

Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., & Thirion, B. (2017, Jan 15). Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *Neuroimage, 145*(Pt B), 166-179. https://doi.org/10.1016/j.neuroimage.2016.10.038

Vincent, L., & Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(6), 583-598. https://doi.org/10.1109/34.87344

Visser, E., Keuken, M. C., Forstmann, B. U., & Jenkinson, M. (2016, Oct 1). Automated segmentation of the substantia nigra, subthalamic nucleus and red nucleus in 7 T data at young and old age. *Neuroimage, 139*, 324-336. https://doi.org/10.1016/j.neuroimage.2016.06.039

Wan, W., Jin, L., Wang, Z., Wang, L., Fei, G., Ye, F., . . . Zhong, C. (2017). Iron Deposition Leads to Neuronal alpha-Synuclein Pathology by Inducing Autophagy Dysfunction. *Front Neurol, 8*, 1. https://doi.org/10.3389/fneur.2017.00001

Wang, J. Y., Zhuang, Q. Q., Zhu, L. B., Zhu, H., Li, T., Li, R., . . . Zhu, J. H. (2016, Nov 9). Meta-analysis of brain iron levels of Parkinson's disease patients determined by postmortem and MRI measurements. *Sci Rep, 6*, 36669. https://doi.org/10.1038/srep36669

Wang, S., Lou, M., Liu, T., Cui, D., Chen, X., & Wang, Y. (2013, Aug). Hematoma volume measurement in gradient echo MRI using quantitative susceptibility mapping. *Stroke, 44*(8), 2315-2317. https://doi.org/10.1161/strokeaha.113.001638

Wang, Y., Spincemaille, P., Liu, Z., Dimov, A., Deh, K., Li, J., . . . Prince, M. R. (2017, Oct). Clinical quantitative susceptibility mapping (QSM): Biometal imaging and its emerging roles in patient care. *J Magn Reson Imaging, 46*(4), 951-971. https://doi.org/10.1002/jmri.25693

Ward, R. J., Zucca, F. A., Duyn, J. H., Crichton, R. R., & Zecca, L. (2014, Oct). The role of iron in brain ageing and neurodegenerative disorders. *Lancet Neurol, 13*(10), 1045-1060. https://doi.org/10.1016/S1474-4422(14)70117-6

Watts, R., Liston, C., Niogi, S., & Uluğ, A. M. (2003). Fiber tracking using magnetic resonance diffusion tensor imaging and its applications to human brain development. *Ment Retard Dev Disabil Res Rev, 9*(3), 168-177. https://doi.org/10.1002/mrdd.10077

Wei, X., Yan, R., Chen, Z., Weng, R., Liu, X., Gao, H., . . . Wang, Q. (2016, 2016/09/20). Combined Diffusion Tensor Imaging and Arterial Spin Labeling as Markers of Early Parkinson's disease. *Sci Rep, 6*(1), 33762. https://doi.org/10.1038/srep33762

Weisenfeld, N. I., Mewes, A. U. J., & Warfield, S. K. (2006). Highly accurate segmentation of brain tissue and subcortical gray matter from newborn MRI. *Medical Image Computing and Computer-Assisted Intervention - Miccai 2006, Pt 1, 4190*, 199-206.

Wood, K.-L., Myall, D. J., Livingston, L., Melzer, T. R., Pitcher, T. L., MacAskill, M. R., . . . Dalrymple-Alford, J. C. (2016, 2016/01/14). Different PD-MCI criteria and risk of dementia in Parkinson's disease: 4-year longitudinal study. *npj Parkinson's Disease, 2*(1), 15027. https://doi.org/10.1038/npjparkd.2015.27

Xiao, Y., Bailey, L., Chakravarty, M. M., Beriault, S., Sadikot, A. F., Pike, G. B., & Collins, D. L. (2012). *Atlas-Based Segmentation of the Subthalamic Nucleus, Red Nucleus, and Substantia Nigra for Deep Brain Stimulation by Incorporating Multiple MRI Contrasts*, Berlin, Heidelberg.

Xiao, Y., Fonov, V., Bériault, S., Al Subaie, F., Chakravarty, M. M., Sadikot, A. F., . . . Collins, D. L. (2015, Mar). Multi-contrast unbiased MRI atlas of a Parkinson's disease population. *Int J Comput Assist Radiol Surg, 10*(3), 329-341. https://doi.org/10.1007/s11548-014-1068-y

Xiao, Y., Jannin, P., D'Albis, T., Guizard, N., Haegelen, C., Lalys, F., . . . Collins, D. L. (2014, Sep). Investigation of morphometric variability of subthalamic nucleus, red nucleus, and substantia nigra in advanced Parkinson's disease patients using automatic segmentation and PCA-based analysis. *Hum Brain Mapp, 35*(9), 4330-4344. https://doi.org/10.1002/hbm.22478

Yablonskiy, D. A., & Haacke, E. M. (1994, Dec). Theory of NMR signal behavior in magnetically inhomogeneous tissues: the static dephasing regime. *Magnetic Resonance in Medicine, 32*(6), 749-763. https://doi.org/10.1002/mrm.1910320610

Yao, B., Li, T. Q., Gelderen, P., Shmueli, K., de Zwart, J. A., & Duyn, J. H. (2009, Feb 15). Susceptibility contrast in high field MRI of human brain as a function of tissue iron content. *Neuroimage, 44*(4), 1259-1266. https://doi.org/10.1016/j.neuroimage.2008.10.029

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006, Jul 1). User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage, 31*(3), 1116-1128. https://doi.org/10.1016/j.neuroimage.2006.01.015

Zhou, D., Liu, T., Spincemaille, P., & Wang, Y. (2014, Mar). Background field removal by solving the Laplacian boundary value problem. *NMR Biomed, 27*(3), 312-319. https://doi.org/10.1002/nbm.3064

Zortea, M., Flores, E., & Scharcanski, J. (2017, 2017/04/01/). A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images. *Pattern Recognition, 64*, 92-104. https://doi.org/https://doi.org/10.1016/j.patcog.2016.10.031